

INTRODUCTION TO PLASMA PHYSICS

With Space and Laboratory Applications

The emphasis of this text is on basic plasma theory, with applications to both space and laboratory plasmas. All mathematical concepts beyond those normally covered in an advanced calculus course are fully explained.

Topics covered include single particle motions, kinetic theory, magnetohydrodynamics, small amplitude waves in both cold and hot plasmas, non-linear phenomena and collisional effects. Applications include planetary magnetospheres and radiation belts, the confinement and stability of plasmas in fusion devices, the propagation of discontinuities and shock waves in the solar wind, and the analysis of various types of plasma waves and instabilities that can occur in planetary magnetospheres and laboratory plasma devices. This book is structured as a text for a one- or two- semester introductory course in plasma physics at the advanced undergraduate or first-year graduate level. It can also serve as a resource book on the basic principles of plasma physics.

DON GURNETT is a pioneer in the study of waves in space plasmas, and has been the author or co-author of over 470 papers in the field of space plasma physics. He is currently a Carver/James A. Van Allen Professor of Physics at the University of Iowa and has received numerous awards for both his teaching and research. In 1994 he received the Iowa Board of Regents Award for Faculty Excellence, and in 1998 was elected a member of the National Academy of Sciences.

AMITAVA BHATTACHARJEE is a leading theoretical plasma physicist and has published over 170 papers on a wide range of subjects spanning fusion, space, and astrophysical plasma physics. He is currently Paul Professor of Space Science at the Institute for the Study of Earth, Oceans, and Space and the Department of Physics at the University of New Hampshire. He is a Fellow of the American Physical Society and the American Association for the Advancement of Science.

INTRODUCTION TO PLASMA PHYSICS

With Space and Laboratory Applications

D. A. GURNETT and A. BHATTACHARJEE

Department of Physics and Astronomy

The University of Iowa



PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge, CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Cambridge University Press 2005

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2005
Reprinted 2006

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 11/14 pt. System L^AT_EX 2_e [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Gurnett, Donald A.

Introduction to plasma physics : with space and laboratory applications / D.A. Gurnett
and A. Bhattacharjee.

p. cm.

Includes bibliographical references and index.

ISBN 0 521 36483 3 – ISBN 0 521 36730 1 (paperback)

1. Plasma (Ionized gases). 2. Space plasmas. I. Bhattacharjee, A. (Amitava), 1955.
II. Title.

QC718.G87 2004
530.4'4 – dc22 2003069745

ISBN 0 521 36483 3 hardback

ISBN 0 521 36730 1 paperback

Contents

	<i>page</i>	<i>ix</i>
<i>Preface</i>		
1 Introduction	1	
2 Characteristic parameters of a plasma	5	
2.1 Number density and temperature	5	
2.2 Debye length	7	
2.3 Plasma frequency	10	
2.4 Cyclotron frequency	12	
2.5 Collision frequency	13	
2.6 Number of electrons per Debye cube	15	
2.7 The de Broglie wavelength and quantum effects	17	
2.8 Representative plasma parameters	18	
3 Single particle motions	23	
3.1 Motion in a static uniform magnetic field	23	
3.2 Motion in perpendicular electric and magnetic fields	26	
3.3 Gradient and curvature drifts	32	
3.4 Motion in a magnetic mirror field	39	
3.5 Motion in a time varying magnetic field	45	
3.6 Adiabatic invariants	48	
3.7 The Hamiltonian method	60	
3.8 Chaotic orbits	68	
4 Waves in a cold plasma	75	
4.1 Fourier representation of waves	75	
4.2 General form of the dispersion relation	84	
4.3 Waves in a cold uniform unmagnetized plasma	87	
4.4 Waves in a cold uniform magnetized plasma	94	
4.5 Ray paths in inhomogeneous plasmas	127	
5 Kinetic theory and the moment equations	137	
5.1 The distribution function	137	
5.2 The Boltzmann and Vlasov equations	140	
5.3 Solutions based on constants of the motion	144	
5.4 The moment equations	146	
5.5 Electron and ion pressure waves	155	
5.6 Collisional drag force	162	
5.7 Ambipolar diffusion	166	
6 Magnetohydrodynamics	175	
6.1 The basic equations of MHD	175	
6.2 Magnetic pressure	183	
6.3 Magnetic field convection and diffusion	185	
6.4 The energy equation	192	
6.5 Magnetohydrodynamic waves	195	

6.6 Static MHD equilibrium	204
6.7 MHD stability	219
6.8 Magnetic reconnection	240
7 Discontinuities and shock waves	251
7.1 The MHD jump conditions	252
7.2 Classification of discontinuities	255
7.3 Shock waves	258
8 Electrostatic waves in a hot unmagnetized plasma	281
8.1 The Vlasov approach	281
8.2 The Landau approach	290
8.3 The plasma dispersion function	308
8.4 The dispersion relation for a multi-component plasma	311
8.5 Stability	318
9 Waves in a hot magnetized plasma	341
9.1 Linearization of the Vlasov equation	342
9.2 Electrostatic waves	345
9.3 Electromagnetic waves	367
10 Non-linear effects	391
10.1 Quasi-linear theory	391
10.2 Stationary non-linear electrostatic potentials	406
11 Collisional processes	415
11.1 Binary Coulomb collisions	416
11.2 Importance of small-angle collisions	417
11.3 The Fokker–Planck equation	420
11.4 Conductivity of a fully ionized plasma	427
11.5 Collision operator for Maxwellian distributions of electrons and ions	431
<i>Appendix A Symbols</i>	435
<i>Appendix B Vector differential operators</i>	441
<i>Appendix C Vector calculus identities</i>	443
<i>Index</i>	445

Preface

This textbook is intended for a full year introductory course in plasma physics at the senior undergraduate or first-year graduate level. It is based on lecture notes from courses taught by the authors for more than three decades in the Department of Physics and Astronomy at the University of Iowa and the Department of Applied Physics at Columbia University. During these years, plasma physics has grown increasingly interdisciplinary, and there is a growing realization that diverse applications in laboratory, space, and astrophysical plasmas can be viewed from a common perspective. Since the students who take a course in plasma physics often have a wide range of interests, typically involving some combination of laboratory, space, and astrophysical plasmas, a special effort has been made to discuss applications from these areas of research. The emphasis of the book is on physical principles, less so on mathematical sophistication. An effort has been made to show all relevant steps in the derivations, and to match the level of presentation to the knowledge of students at the advanced undergraduate and early graduate level. The main requirements for students taking this course are that they have taken an advanced undergraduate course in electricity and magnetism and that they are knowledgeable about using the basic principles of vector calculus, i.e., gradient, divergence and curl, and the various identities involving these vector operators. Although extensive use is made of complex variables, no special back-ground is required in this subject beyond what is covered in an advanced calculus course. Relatively advanced mathematical concepts that are not typically covered in an undergraduate sequence, such as Fourier transforms, Laplace transforms, the Cauchy integral theorem, and the residue theorem, are discussed in sufficient detail that no additional preparation is required. Although this approach has undoubtedly added to the length of the book, we believe that the material covered provides an effective and self-contained textbook for teaching plasma physics. MKS units are used throughout.

For the preparation of this text we would especially like to thank Kathy Kurth who did the typing and steadfastly stuck with us through the many revisions and additions that occurred over the years. We would also like to thank Joyce Chrisinger and Ann Persoon for their outstanding work preparing the illustrations and proofreading, and Dr. C.-S. Ng and Dr. Z. Ma for checking the accuracy of the equations. Don Gurnett would like to acknowledge the salary support provided by the Carver Foundation during the preparation of this manuscript, and Amitava Bhattacharjee would like to acknowledge the generous support of the Faculty Scholar Program at the University of Iowa and the Peter Paul Chair at the University of New Hampshire.

A plasma is an ionized gas consisting of positively and negatively charged particles with approximately equal charge densities. Plasmas can be produced by heating an ordinary gas to such a high temperature that the random kinetic energy of the molecules exceeds the ionization energy. Collisions then strip some of the electrons from the atoms, forming a mixture of electrons and ions. Because the ionization process starts at a fairly well-defined temperature, usually a few thousand K, a plasma is often referred to as the “fourth” state of matter. Plasmas can also be produced by exposing an ordinary gas to energetic photons, such as ultraviolet light or X-rays. The steady-state ionization density depends on a balance between ionization and recombination. In order to maintain a high degree of ionization, either the ionization source must be very strong, or the plasma must be very tenuous so that the recombination rate is low.

The definition of a plasma requires that any deviation from charge neutrality must be very small. For simplicity, unless stated otherwise, we will assume that the ions are singly charged. The charge neutrality condition is then equivalent to requiring that the electron and ion number densities be approximately the same. In the absence of a loss mechanism, the overall charge neutrality assumption is usually satisfied because all ionization processes produce equal amounts of positive and negative charge. However, deviations from *local* charge neutrality can occur. Usually these deviations are small, since as soon as a charge imbalance develops, large electric fields are produced that act to restore charge neutrality. Systems that display large deviations from charge neutrality, such as vacuum tubes and various electronic devices, are not plasmas, even though some aspects of their physics are similar.

To be a plasma the charged particles must be in an unbound gaseous state. This requirement can be made more specific by requiring that the random kinetic energy be much greater than the average electrostatic energy, and is imposed to provide a distinction between a plasma, in which the particles move relatively freely, and condensed matter, such as metals, where electrostatic forces play a dominant role. In a plasma long-range electrical forces are much more important than short-range forces. Because many particles “feel” the same long-range forces, a plasma is dominated by “collective” motions involving correlated movements of large numbers of particles rather than uncorrelated interactions between neighboring particles. Long-range forces lead to many complex effects that do not occur in an ordinary gas.

Plasmas can be divided into two broad categories: natural and man-made. It is an interesting fact that most of the material in the visible universe, as much as 99% according to some estimates, is in the plasma state. This includes the Sun, most stars, and a significant fraction of the interstellar medium. Thus, plasmas play a major role in the universe.

Plasma physics is relevant to the formation of planetary radiation belts, the development of sunspots and solar flares, the acceleration of high velocity winds that flow outward from the Sun and other stars, the generation of radio emissions from the Sun and other astrophysical objects, and the acceleration of cosmic rays.

In the Earth's atmosphere, the low temperatures and high pressures that are commonly present are not favorable for the formation of plasmas except under unusual conditions. Probably the most common plasma phenomenon encountered in the Earth's atmosphere is lightning. In a lightning discharge the atmospheric gas is ionized and heated to a very high temperature by the electrical currents that are present in the discharge. Because of the high recombination rate the resulting plasma exists for only a small fraction of a second. Less common is ball lightning, which consists of a small ball of hot luminous plasma that lasts for up to tens of seconds. Another terrestrial plasma phenomenon, readily observable at high latitudes, is the aurora, which is produced by energetic electrons and ions striking the atmosphere at altitudes of 80 to 100 km. At higher altitudes, from one hundred to several hundred km, the Earth is surrounded by a dense plasma called the ionosphere. The ionospheric plasma is produced by ultraviolet radiation from the Sun, and also exists on the nightside of the Earth because the recombination rate is very low at high altitudes. The ionosphere plays an important role in radio communication by acting as a reflector for low-frequency radio waves. At even higher altitudes, the Earth is surrounded by a region of magnetized plasma called the magnetosphere. Planetary magnetospheres have now been observed at all the magnetized planets and exhibit many of the plasma processes that are believed to occur at magnetized astronomical objects such as neutron stars.

Numerous applications of basic plasma physics can be found in man-made devices. One of the most important of these is the attempt to achieve controlled thermonuclear fusion. Because fusion requires temperatures of 10^7 K, or more, to overcome the Coulomb repulsion between nuclei, controlled fusion necessarily involves very high temperatures. Since a fusion plasma would be quickly cooled by the walls of any ordinary container, considerable effort has gone into attempts to contain plasmas by magnetic fields, using a so-called "magnetic bottle." Although the principles of such magnetic confinement may appear at first glance to be straightforward, attempts to achieve controlled fusion using magnetic confinement have been complicated by collective effects that develop when large numbers of particles are introduced into the machine. The effort to find a technologically and economically attractive configuration for confining a dense, hot plasma remains one of the main challenges of fusion research. Besides fusion, numerous other devices involving plasmas also exist. Fluorescent lights and various other devices involving plasma discharges, such as electric arc welders and plasma etching machines, are in common daily use. More advanced devices include magnetohydrodynamic generators for producing electricity from high-temperature gas jets, ion engines for spacecraft propulsion, various surface treatment processes that involve the injection of ions into metal surfaces, and high-frequency electronic devices such as traveling wave tubes and magnetrons.

The purpose of this book is to provide the basic principles needed to analyze a broad range of plasma phenomena. Since both natural and man-made plasmas are of potential interest, a special effort has been made in this book to provide examples from both space and laboratory applications.

Characteristic parameters of a plasma

Before starting with a detailed discussion of the processes that occur in a plasma, it is useful to identify certain fundamental parameters that are relevant to the description of essentially all plasma phenomena.

2.1 Number density and temperature

In an ordinary material there are usually three parameters, pressure, density and temperature, that must be specified to determine the state of the material, any two of which can be selected as the independent variables. A plasma almost always involves considerably more parameters. For a plasma consisting of electrons and various types of ions, it is necessary to define a number density for each species, denoted by n_s where the subscript s stands for the s th species. Since the electrons and ions respond differently to electromagnetic forces, the number densities of the various species must be regarded as independent variables. In general, a plasma cannot be characterized by a single density.

The temperature of particles of type s is directly proportional to their average random kinetic energy. In thermal equilibrium the distribution of velocities of particles of type s is given by the Maxwellian distribution

$$f_s(\mathbf{v}) = n_s \left(\frac{m_s}{2\pi\kappa T_s} \right)^{3/2} e^{-\frac{m_s v^2}{2\kappa T_s}}, \quad (2.1.1)$$

where $f_s(\mathbf{v})$ is the distribution function, \mathbf{v} is the velocity, m_s is the mass of the particles, κ is Boltzmann's constant, and T_s is the temperature. The distribution function is normalized such that $f_s(\mathbf{v})$ integrated over all velocities gives the number density of particles of type s ,

$$\int_{-\infty}^{\infty} f_s(\mathbf{v}) dv_x dv_y dv_z = n_s. \quad (2.1.2)$$

A plot of the Maxwellian distribution as a function of v_x is shown in Figure 2.1 (for $v_y = v_z = 0$). It is a relatively simple matter to show that the root-mean-square velocity is given by $\sqrt{3}C_s$, where

$$C_s = \sqrt{\frac{\kappa T_s}{m_s}}. \quad (2.1.3)$$

Hereafter, C_s will be referred to as the thermal speed. The average kinetic energy is given by

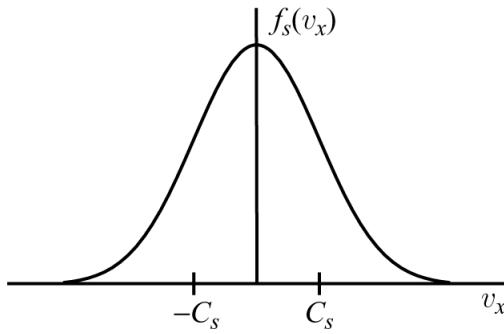


Fig. 2.1 The Maxwellian velocity distribution.

$$\left\langle \frac{1}{2}m_s v^2 \right\rangle = \frac{3}{2}\kappa T_s, \quad (2.1.4)$$

where the angle brackets indicate an average. The above equation shows that the temperature is directly proportional to the average kinetic energy of the particles.

According to a general principle of statistical mechanics called the H-theorem, the Maxwellian distribution is the unique distribution function that arises when a gas is in thermal equilibrium [Huang, 1963]. For a plasma in thermal equilibrium, not only should the distribution function for each species be a Maxwellian but the temperature of all species must be equal. However, because collisions occur very infrequently in a tenuous plasma, the approach to thermal equilibrium is often very slow. Therefore, non-equilibrium effects are quite common in plasmas. Since the electron and ion masses are very different, the rate of energy transfer between electrons and ions is much slower than between electrons or between ions. Therefore, when a plasma is heated, substantial temperature differences often develop between the electrons and ions. Non-equilibrium distributions also occur when an electron beam or an ion beam is injected into a plasma. Under these circumstances the velocity distribution function of the beam usually cannot be represented by a Maxwellian distribution. Such non-thermal distributions produce many interesting effects that will be discussed later.

2.2 Debye length

All plasmas are characterized by a fundamental length scale determined by the temperature and number density of the charged particles. To demonstrate the existence of this length scale consider what happens when a negative test charge Q is placed in an otherwise homogeneous plasma. Immediately after the charge is introduced, the electrons are repelled and the ions are attracted. Very quickly, the resulting displacement of the electrons and ions produces a polarization charge that acts to shield the plasma from the test charge. This shielding effect is called Debye shielding, after Debye and Hückel [1923] who first studied

the effect in dielectric fluids. The characteristic length over which shielding occurs is called the Debye length.

To obtain an expression for the Debye length, it is useful to consider a homogeneous plasma of electrons of number density n_e and temperature T_e , and a fixed background of positive ions of number density n_0 . After the negative test charge Q has been inserted and equilibrium has been established, the electrostatic potential Φ is given by Poisson's equation

$$\nabla^2 \Phi = -\frac{\rho_q}{\epsilon_0} = -\frac{e}{\epsilon_0}(n_0 - n_e), \quad (2.2.1)$$

where ρ_q is the charge density, ϵ_0 is the permittivity of free space and e is the electronic charge. To obtain a solution for the electrostatic potential, it is necessary to specify the electron density as a function of the electrostatic potential. We assume that at infinity, where $\Phi = 0$, the electrons have a Maxwellian velocity distribution with a number density n_0 . From general principles of kinetic theory it can be shown that the velocity distribution function for the electrons is given by

$$f_e(v) = n_0 \left(\frac{m_e}{2\pi\kappa T_e} \right)^{3/2} e^{-\frac{(\frac{1}{2}m_e v^2 + q\Phi)}{\kappa T_e}}, \quad (2.2.2)$$

where $q = -e$. This equation is like the Maxwellian distribution discussed previously, but has an additional factor $\exp[-q\Phi/\kappa T]$. This factor comes from a principle of statistical mechanics that states the number of particles with velocity v is proportional to $\exp[-W/\kappa T]$, where W is the total energy [Huang, 1963]. The total energy is given by the sum of the kinetic energy and the potential energy, $W = (1/2)m_e v^2 + q\Phi$. By integrating the distribution function over velocity space, it is easy to show that the electron density is given by

$$n_e = n_0 e^{\frac{e\Phi}{\kappa T_e}}. \quad (2.2.3)$$

Substituting the above expression into Poisson's equation (2.2.1), one obtains the non-linear differential equation

$$\nabla^2 \Phi = -\frac{n_0 e}{\epsilon_0} \left(1 - e^{\frac{e\Phi}{\kappa T_e}} \right). \quad (2.2.4)$$

This differential equation can be solved analytically if we assume that $e\Phi/\kappa T_e \ll 1$. Expanding the exponential in a Taylor series and keeping only the first-order term, one obtains the linear differential equation

$$\nabla^2 \Phi = \frac{n_0 e^2}{\epsilon_0 \kappa T_e} \Phi. \quad (2.2.5)$$

Since the plasma is isotropic, the electrostatic potential can be assumed to be spherically symmetric. The above equation then simplifies to

$$\frac{\partial^2}{\partial r^2} (r\Phi) - \frac{n_0 e^2}{\epsilon_0 \kappa T_e} (r\Phi) = 0, \quad (2.2.6)$$

which has the general solution

$$\Phi = \frac{A}{r} e^{-r/\lambda_D}, \quad (2.2.7)$$

where r is the radius and A is a constant. The factor λ_D is the Debye length and is easily shown to be given by

$$\lambda_D^2 = \frac{\epsilon_0 \kappa T_e}{n_0 e^2}. \quad (2.2.8)$$

The constant A is determined by requiring that the solution reduce to the Coulomb potential as the radius goes to zero. The complete solution is then given by

$$\Phi = \frac{1}{4\pi\epsilon_0} \frac{Q}{r} e^{-r/\lambda_D}, \quad (2.2.9)$$

and is called the Debye-Hückel potential. A plot of the Debye-Hückel potential (for negative Q) is shown in Figure 2.2. As can be seen, the potential decays exponentially, with a length scale given by the Debye length, λ_D . A simple, practical formula for the Debye length is

$$\lambda_D = 6.9 \sqrt{T_e / n_0} \text{ cm}, \quad (2.2.10)$$

where T_e is in K and n_0 is in cm^{-3} .

The derivation of the Debye length given above is deceptively simple and hides some subtleties inherent in the concept, especially in collisionless plasmas for which the assumption of a Maxwellian distribution (2.2.2) is open to question. For instance, consider the following interesting paradox involving the role of ions. If the ions were mobile, then by a simple extension of the treatment used for the electrons it would appear that the ion number density should be given by

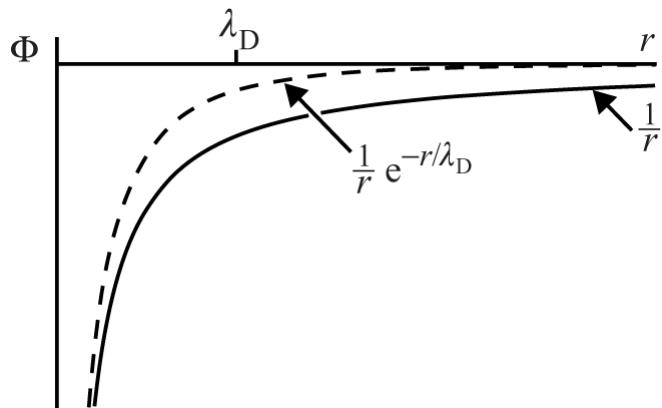


Fig. 2.2 A comparison of the Debye-Hückel potential (dashed line) with the Coulomb potential (solid line) for a negative test charge.

$$n_i = n_0 e^{-\frac{e\Phi}{kT_i}}. \quad (2.2.11)$$

However, this equation does not provide a correct representation of the ion density. This is because the ions are accelerated by the negative charge Q , and particle flux conservation dictates that as the ion velocity increases, the ion density should decrease. This suggests that the mobility of the ions actually decreases the positive charge density, in contrast with equation (2.2.11) which implies that the ion density should increase as r decreases. This effect is called *anti-shielding*, since it decreases the charge density in the shielding region. The resolution of this paradox requires a more sophisticated understanding of distribution functions than we have at the moment, and is postponed to Chapter ??.

2.2.1 Plasma sheaths

When an object of finite size is placed in a plasma with approximately equal electron and ion temperatures, it acquires a net negative charge because the electron thermal speed, $C_e = \sqrt{\kappa T_e/m_e}$, is much greater than the ion thermal speed, $C_i = \sqrt{\kappa T_i/m_i}$, which causes more electrons to hit the object than ions. As the object charges negatively, the electrons start to be repelled, just as when a negative test charge is introduced into the plasma. Equilibrium occurs when the electron current collected by the object balances the incident ion current. An electrically polarized region is thereby formed around the object. This polarized region is called a plasma sheath, or sometimes a positive ion sheath, because the electrons are largely excluded from the sheath. The exact form of the electrostatic potential distribution is a complicated boundary value problem and can only be solved analytically for certain simple geometries such as a sphere, a cylinder, or a planar surface. If the radius of curvature is much larger than the Debye length, so that the surface can be regarded as locally planar, then the potential decays exponentially with a characteristic length scale given by the Debye length. In these simple cases it is easy to show, by equating the incident electron and ion currents, that the equilibrium potential of the surface is given to a good approximation by

$$V = -\frac{\kappa T_e}{2e} \left[\ln\left(\frac{m_i}{m_e}\right) + \ln\left(\frac{T_e}{T_i}\right) \right]. \quad (2.2.12)$$

Note that for a given ion-to-electron mass ratio the equilibrium potential is controlled entirely by the electron temperature. Because of the weak logarithmic dependence, this potential is typically only a few times the electron thermal energy. For a proton–electron plasma with equal electron and ion temperatures $V = -3.75\kappa T/e$.

If the object is exposed to ultraviolet radiation, as in the case of a spacecraft exposed to sunlight, then the emitted photoelectron current must be added to the equilibrium current balance condition. Under these conditions the object can charge to a positive potential if the photoelectron flux exceeds the incident electron flux. Modifications to the equilibrium potential can also occur if secondary electrons are produced by energetic particles striking the surface.

2.3 Plasma frequency

If the electrons in a uniform, homogeneous plasma are displaced from their equilibrium position, an electric field arises because of charge separation. This electric field produces a restoring force on the displaced electrons. Since the magnitude of the charge imbalance is directly proportional to the displacement, the restoring force is given by Hooke's law, $F = -k\Delta x$, where Δx is the displacement and k is the effective "spring constant." Since the electrons have inertia, the system behaves as a harmonic oscillator. The resulting oscillations are called electron plasma oscillations or Langmuir oscillations, after Tonks and Langmuir [1929] who first discovered these oscillations.

To compute the oscillation frequency, let us assume that the plasma consists of a uniform slab of electrons of number density n_0 and a fixed background of positive ions of the same density. Suppose we now displace the slab of electrons to the right by a small distance Δx , as shown in Figure 2.3. The slab can be divided into three regions. Region 1 has a net positive charge, region 2 has no net charge, and region 3 has a net negative charge. The electric field in region 2 can be computed using Gauss' law and is given by

$$E = \frac{n_0 e \Delta x}{\epsilon_0}. \quad (2.3.1)$$

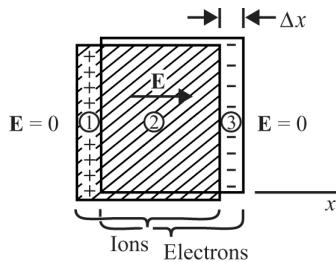


Fig. 2.3 A simple slab model that illustrates electron plasma oscillations.

If the slab of electrons is then released, the equation of motion for the electrons is given by

$$m_e \frac{d^2 \Delta x}{dt^2} = (-e)E = -\frac{n_0 e^2}{\epsilon_0} \Delta x, \quad (2.3.2)$$

which simplifies to

$$\frac{d^2 \Delta x}{dt^2} + \left(\frac{n_0 e^2}{\epsilon_0 m_e} \right) \Delta x = 0. \quad (2.3.3)$$

The above equation is just the harmonic oscillator equation. The oscillation frequency ω_{pe} is determined by the term in parentheses

$$\omega_{pe}^2 = \frac{n_0 e^2}{\epsilon_0 m_e}, \quad (2.3.4)$$

and is called the electron plasma frequency. A simple formula for the electron plasma frequency (in hertz) is given by

$$f_{pe} = 8980 \sqrt{n_0} \text{Hz}, \quad (2.3.5)$$

where the number density n_0 is in electrons cm^{-3} . Note that the electron plasma frequency is determined solely by the number density of the electrons.

If a plasma contains several species, it is customary to define a plasma frequency for each species according to the equation

$$\omega_{ps}^2 = \frac{n_s e_s^2}{\epsilon_0 m_s}, \quad (2.3.6)$$

where e_s , m_s , and n_s are the charge, mass, and number density of that species. It is easily verified that the plasma frequency, the Debye length, and the thermal speed are related by the formula

$$\omega_{ps} \lambda_{Ds} = C_s. \quad (2.3.7)$$

It should *not* be inferred that in a multi-species plasma, each species oscillates independently at its own plasma frequency. If the ions are allowed to move in the simple slab model discussed above, it is easy to show that the oscillation frequency is given by

$$\omega_p = \sqrt{\omega_{pe}^2 + \omega_{pi}^2}, \quad (2.3.8)$$

which is approximately ω_{pe} , since $\omega_{pi} \ll \omega_{pe}$.

2.4 Cyclotron frequency

It is well known that if a charged particle of mass m and charge q is injected into a uniform static magnetic field, it moves in a uniform circular motion around the magnetic field at a characteristic frequency called the cyclotron frequency, given by

$$\omega_c = \frac{|q|B}{m}, \quad (2.4.1)$$

where B is the magnetic field strength. A simple formula for the electron cyclotron frequency (in hertz) is given by

$$f_{ce} = 28B \text{Hz}, \quad (2.4.2)$$

where B is in nT. We adopt here the standard convention that the cyclotron frequency is a positive quantity, independent of the sign of the charge. However, if a plasma involves several species, it is often convenient to write the cyclotron frequency, which occurs in the

equations of motion, in a form that includes the sign of the charge. In this case, to simplify notation, the sign of the charge is included in the cyclotron frequency, with ω_{cs} defined as

$$\omega_{cs} = \frac{e_s B}{m_s}, \quad (2.4.3)$$

where e_s is the charge for that species (including sign). Since the ions normally encountered in plasmas are positively charged, in this book we always assume that the ion cyclotron frequency, ω_{ci} , is positive. For electrons, the sign of the cyclotron frequency is negative. However, when only electrons are being considered, it will be our convention to use ω_c for the magnitude of the electron cyclotron frequency, as in equation (2.4.1), and incorporate the minus sign into the equations involved in the analysis.

2.5 Collision frequency

To describe the effects of collisions between particles in a plasma, it is convenient to define the collision frequency, ν_{rs} , as the average rate at which particles of type r collide with particles of type s . Two types of collisions can be defined: those between charged particles and neutral particles, and those between charged particles. The physics involved in these two types of collisions is quite different, and is discussed separately.

2.5.1 Collisions between charged and neutral particles

For collisions between charged particles (of type s) and neutral particles (of type n), the interaction force has a very short range and the scattering process is similar to the scattering produced by hard spheres. If we assume that the neutrals are immobile hard spheres, it is easy to show that the average collision frequency is given by

$$\nu_{ns} = n_n C_s \sigma_n, \quad (2.5.1)$$

where n_n is the number density of the neutral gas, σ_n is the collision cross-section with neutral atoms, and C_s is the thermal speed of charged particles of type s , defined by equation (2.1.3). Note that the collision frequency increases as the temperature increases.

The effect of neutral collisions on a plasma is determined by the collision frequency and the time scale τ of the plasma process of interest. If $v\tau$ is much less than one, the effect of neutral collisions on the plasma is small. In this weakly collisional regime, the main effect is a weak damping of disturbances and waves, with a slow rate of energy transfer to the neutrals. To achieve this regime, the number density of neutral atoms (and hence the neutral gas pressure) must be reduced to a low level.

If, on the other hand, $v\tau$ is much greater than one, the plasma is said to be collision dominated. In this case, the motion of the ionized component is almost completely controlled by the motion of the neutral gas, and plasma processes are of secondary importance.

2.5.2 Collisions between charged particles

Because the Coulomb force is long range, collisions between charged particles are quantitatively different than collisions between charged particles and neutral particles. For such collisions, it is useful to define the differential scattering cross-section, $\sigma_C(\chi)$, by

$$\sigma_C(\chi)d\Omega = \frac{\text{number of particles scattered into } d\Omega}{\text{incident beam intensity}}, \quad (2.5.2)$$

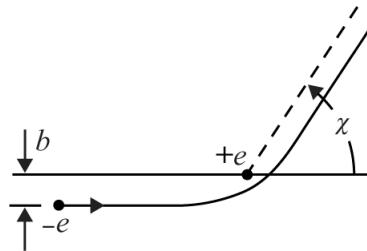


Fig. 2.4 A typical trajectory for Coulomb scattering.

where χ is the scattering angle relative to the incident beam, and $d\Omega = 2\pi \sin \chi d\chi$ is the differential solid angle. For an electron of mass m_e and charge $-e$ incident on a much heavier ion of charge $+e$ and mass m_i , where $m_i \gg m_e$, the differential scattering cross-section is given by the Rutherford formula

$$\sigma_C(\chi) = \frac{1}{4} \left(\frac{e^2}{4\pi\epsilon_0 m_e v^2} \right)^2 \frac{1}{\sin^4(\chi/2)}. \quad (2.5.3)$$

Simple principles of classical mechanics show that small scattering angles are associated with large impact parameters, b , and vice versa. The scattering geometry is illustrated in Figure 2.4. In contrast to scattering by neutral particles, Coulomb scattering is highly anisotropic and has a strong peak at small angles, as can be readily seen from the inverse $\sin^4(\chi/2)$ dependence. Furthermore, the total scattering cross-section, obtained by integrating expression (2.5.3) over a 4π solid angle,

$$\sigma_T = \int_0^\pi \sigma_C(\chi) 2\pi \sin \chi d\chi \quad (2.5.4)$$

is infinite because the integral diverges at the lower limit, which corresponds to large impact parameters. In a plasma, this divergence does not occur because the Coulomb force is strongly reduced by Debye shielding for impact parameters greater than the Debye length. Thus, the Debye length sets a lower limit on the scattering angle (or equivalently, an upper limit on the impact parameter), and eliminates the divergence in equation (2.5.4).

It is interesting to note that the Coulomb collision cross-section varies as $1/v^4$. If we interpret v as the thermal speed of the incident particles, it follows then that the collision frequency $\nu = n v \sigma_C$ varies as $1/v^3$, or as $1/T^{3/2}$. Thus, the frequency for collisions with

charged particles decreases as the temperature increases, which is just the opposite of collisions with neutral particles (see Section 2.5.1). In the absence of neutral collisions, a nearly collisionless regime is attained when the temperature is very large. Later in Chapter ??, we show that the effective collision frequency for electrons colliding with ions is given by

$$\nu_{ei} = \frac{n_0 e^4}{32\pi^{1/2} \epsilon_0^2 m_e^{1/2} (2\kappa T_e)^{3/2}} \ln(12\pi N_D), \quad (2.5.5)$$

where $N_D = n_0 \lambda_D^3$ is the number of electrons per Debye cube. The logarithmic term implies that the numerical value of the collision frequency is weakly dependent on N_D .

2.6 Number of electrons per Debye cube

In the previous section, we encountered the dimensionless parameter $N_D = n_0 \lambda_D^3$, which is the number of electrons per Debye cube. This parameter plays a fundamental role in the description of a plasma, and is proportional to a quantity called the plasma parameter (see Chapter ??). The importance of the number of electrons per Debye cube can be illustrated from several points of view.

2.6.1 Macroscopic averages

An implicit assumption in our analysis of a plasma is that macroscopic averages are meaningful. For example, consider the electron density, defined as the average number of electrons per unit volume. In order that the electron density be a smooth function of space, it is necessary that we calculate the average over a volume element that includes many particles.

To consider the effect the number of particles has on this averaging process, let us start with an arbitrary collection of particles of charge e and mass m and divide each particle into M equal parts. Upon division, the charge and mass of each particle become $e' = e/M$ and $m' = m/M$, while the number density becomes $n' = nM$. Since the temperature is proportional to the mass, $(1/2)mv^2 = (3/2)\kappa T$, the new temperature becomes $T' = T/M$. As we take the limit $M \rightarrow \infty$, the discrete nature of the particles is suppressed and the system tends to a continuous fluid. It is easy to show that the fundamental length and time scales of the plasma are unchanged by this fluidization process: i.e., $\lambda'_D = \lambda_D$, $\omega'_p = \omega_p$, and $\omega'_c = \omega_c$. On the other hand, in this limiting process, the number of electrons per Debye cube increases without bound because $N'_D = N_D M$. The parameter N_D , therefore, characterizes the extent to which we can treat the plasma as a fluid: a plasma can be considered a continuous fluid only if N_D is very large (i.e., much greater than one).

As another illustration of the role of N_D in macroscopic averages, we return to the discussion of Debye shielding in Section 2.2. In the derivation of the Debye shielding formula (2.2.7), we assumed that $e\Phi/\kappa T_e \ll 1$ in order to obtain a tractable expression for the charge density. At the time, no justification was given for this assumption, which is clearly

violated at a sufficiently small radius, since the Coulomb potential varies as $1/r$. Since the charge density involves a macroscopic average, we can ask how small the radius r can possibly be and still have a meaningful average number density. As a rough approximation, we take this minimum radius to correspond to a volume which contains only one electron, i.e., $r_{\min} = 1/n_e^{1/3}$. The Coulomb potential at this radius is approximately

$$\Phi = \frac{1}{4\pi\epsilon_0} \frac{e}{r_{\min}}. \quad (2.6.1)$$

The $e\Phi/\kappa T_e$ term is then given by

$$\frac{e\Phi}{\kappa T_e} = \frac{1}{4\pi\epsilon_0} \frac{e^2}{kT_e} n_e^{1/3}. \quad (2.6.2)$$

With a slight rearrangement of the terms, it is easy to show that the above equation can be rewritten as

$$\frac{e\Phi}{\kappa T_e} = \frac{1}{4\pi} \frac{1}{N_D^{2/3}}, \quad (2.6.3)$$

where $N_D = n_0\lambda_D^3$. One can see that the assumption $e\Phi/\kappa T_e \ll 1$ is satisfied only if the number of electrons per Debye cube is much greater than one ($N_D \gg 1$).

2.6.2 Ratio of kinetic energy to potential energy

We next show that the number of electrons in a Debye cube is closely related to the ratio of the average kinetic energy to the average electrostatic potential energy. This ratio can be estimated by noting that the average kinetic energy of a particle is $(3/2)\kappa T$ and the average electrostatic energy is approximately $e^2/4\pi\epsilon_0\langle r \rangle$, where $\langle r \rangle$ is the average separation distance between neighboring particles. The ratio of these two quantities gives

$$\frac{\text{kinetic energy}}{\text{potential energy}} = \frac{3\kappa T 4\pi\epsilon_0 \langle r \rangle}{2e^2}. \quad (2.6.4)$$

Since $\langle r \rangle \simeq 1/n^{1/3}$, the above equation simplifies to

$$\frac{\text{kinetic energy}}{\text{potential energy}} = 6\pi N_D^{2/3}. \quad (2.6.5)$$

If $N_D \gg 1$, one can see that the average kinetic energy greatly exceeds the average electrostatic potential energy. Under these conditions, the plasma is dominated by small-angle collisions. Large-angle collisions caused by the close approach of pairs of particles are extremely rare. As N_D decreases, electrostatic interactions become more important and large-angle collisions become more frequent. As N_D drops below one, electrostatic interactions become dominant and the motions are strongly controlled by the electric fields in the near vicinity of the individual charges. Electrons become trapped in the potential wells of the ions and the ions organize their average position so as to minimize the average

electrostatic energy, leading to lattice-like structures. Systems of this type are said to be strongly coupled, and are characteristic of solids.

2.6.3 Ratio of collective to discrete interactions

Because of the long-range nature of the Coulomb force, a plasma is dominated by large-scale collective motions in which many particles respond in unison to a given electric field. Electron plasma oscillations are an example of such a collective effect. Discrete collisional interactions, on the other hand, tend to inhibit such large-scale collective motions since the collisional forces only act on individual pairs of particles. The importance of collective interactions relative to discrete interactions can be characterized by taking the ratio of the electron plasma frequency to the electron collision frequency. Using equations (2.3.4) and (2.5.5), this ratio can be written entirely as a function of N_D ,

$$\frac{\omega_{pe}}{\nu_{ei}} = \sqrt{\frac{\pi}{2}} \frac{128N_D}{\ln(12\pi N_D)}. \quad (2.6.6)$$

If $N_D \gg 1$ the electron plasma frequency is much greater than the electron-ion collision frequency. This means that collective oscillations, characterized by ω_{pe} , are much more important than discrete-particle effects, which are characterized by ν_{ei} . On the other hand, if $N_D \ll 1$, plasma oscillations are rapidly destroyed by collisions between individual particles. This situation occurs in solids, such as metals, where the oscillations at the electron plasma frequency are strongly damped by collisional interactions.

2.7 The de Broglie wavelength and quantum effects

Quantum effects become important if the typical distance separating particles in a plasma becomes comparable or less than the de Broglie wavelength of the particles. Since the de Broglie wavelength of an electron is much larger than that of an ion, quantum effects occur for electrons before they do for ions. Therefore, to determine the boundary between classical and quantum regimes, it is sufficient to consider only electrons. Heisenberg's uncertainty principle states that

$$\langle \Delta p \rangle \langle \Delta x \rangle \geq \frac{\hbar}{2}, \quad (2.7.1)$$

where $\langle \Delta p \rangle$ and $\langle \Delta x \rangle$ are the uncertainty in the momentum and position, respectively, and \hbar is Planck's constant divided by 2π [Merzbacher, 1998]. For a classical description, the largest possible $\langle \Delta p \rangle$ occurs when the uncertainty in the momentum is equal to the root-mean-square (r.m.s.) momentum, which using equation (2.1.4) is given by

$$\langle \Delta p \rangle = \langle (m_e v)^2 \rangle^{1/2} = (3m_e \kappa T_e)^{1/2}. \quad (2.7.2)$$

Quantum effects first become important when the uncertainty in the position $\langle \Delta x \rangle$ is comparable to the average distance to the nearest electron, which is approximately

$$\langle \Delta x \rangle = 1/n_e^{1/3}. \quad (2.7.3)$$

Combining conditions (2.7.2) and (2.7.3) with the uncertainty principle (2.7.1), gives the following condition for the validity of the classical description:

$$\frac{(3m_e\kappa T_e)^{1/2}}{n_e^{1/3}} \gg \hbar. \quad (2.7.4)$$

The above inequality defines a boundary in $n_e - T_e$ space that separates the classical from the quantum regime. The equation defining this boundary is given by

$$\frac{T_e^{3/2}}{n_e} = \frac{(\hbar/2)^3}{(3m_e\kappa)^{3/2}}. \quad (2.7.5)$$

It is interesting to note that quantum effects become more important as the temperature decreases. This occurs because the de Broglie wavelength increases as the electron velocity (temperature) decreases. The numerical value of the factor on the right-hand side of the above equation is $6.33 \times 10^{-19} \text{ K}^{3/2} \text{ cm}^{-3}$, which is very small. So only very moderate temperatures are required for the classical description to be valid. On the other hand, the conditions for the quantum regime are met at very high densities, as in metals and various astrophysical objects such as white dwarfs.

2.8 Representative plasma parameters

The basic parameters described above can be summarized by means of the simple diagram in Figure 2.5, which shows the range of temperatures and number densities for various plasmas. For convenience, we also list the temperatures in electron volts, $\kappa T_e/e$, and the number densities in terms of the electron plasma frequency. Lines of constant Debye length and constant N_D are also shown. The so-called quantum degeneracy line separates the classical regime above from the quantum regime below. This book is entirely concerned with plasmas in the classical regime, with $N_D \gg 1$.

Problems

- 2.1. Consider a plasma consisting of electrons and ions of equal number densities, $n = n_e = ni$, and temperatures, T_e and T_i . Assume that the electron density is determined by a term of the form $\exp[e\Phi/\kappa T_e]$ and that the ion density is determined by a term of the form $\exp[-e\Phi/\kappa T_i]$. Show that the Debye length, λ_D , is then given by

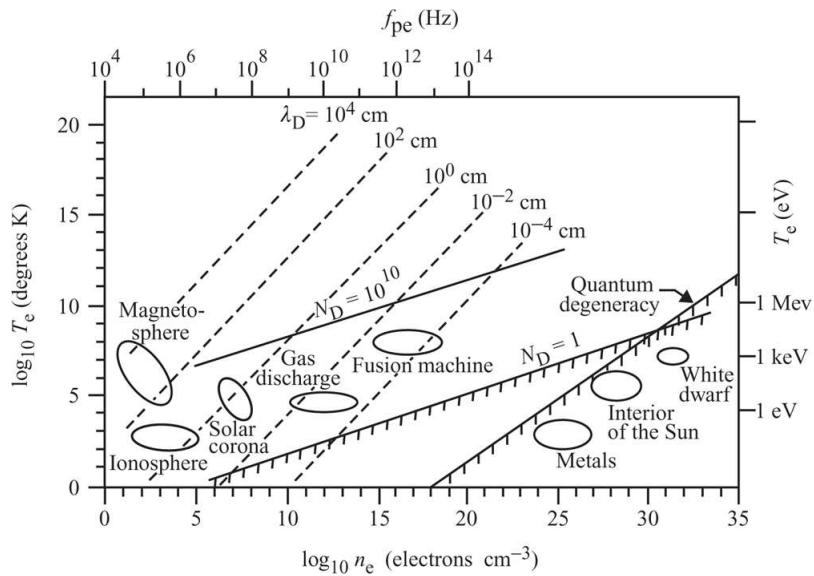


Fig. 2.5 The temperature and density of various plasmas.

$$\left(\frac{1}{\lambda_D}\right)^2 = \left(\frac{1}{\lambda_{De}}\right)^2 + \left(\frac{1}{\lambda_{Di}}\right)^2.$$

How do the electron and ion Debye length terms, λ_{De} and λ_{Di} , compare if $T_e = T_i$? Can you find any reason to criticize the basic model used in this calculation?

- 2.2. Consider two superposed slabs of electrons of mass m_e and ions of mass m_i . Using the same type of analysis performed in the text, show that the natural oscillation frequency is

$$\omega_p^2 = \omega_{pe}^2 + \omega_{pi}^2$$

where ω_{pe} and ω_{pi} are the electron and ion plasma frequencies (assume that the ions have charge e).

- 2.3. Demonstrate the following relationship between the electron plasma frequency, ω_{pe} , the electron thermal speed, C_e , and the Debye length, λ_{De} :

$$C_e = \omega_{pe} \lambda_{De}$$

or equivalently, defining $k_D = 1/\lambda_D$,

$$k_{De} C_e = \omega_{pe}.$$

- 2.4. A Maxwellian velocity distribution has the form

$$f(\mathbf{v}) = n_0 \left(\frac{m}{2\pi\kappa T} \right)^{3/2} \exp \left[-\frac{mv^2}{2\kappa T} \right],$$

where $v^2 = v_x^2 + v_y^2 + v_z^2$.

(a) Show that

$$\int_{-\infty}^{\infty} f(\mathbf{v}) dv_x dv_y dv_z = n_0.$$

(b) Show that

$$\left\langle \frac{1}{2}mv_x^2 \right\rangle = \frac{1}{2}\kappa T.$$

(c) Show that

$$\left\langle \frac{1}{2}mv^2 \right\rangle = \frac{3}{2}\kappa T.$$

- 2.5. An infinite plate is immersed in a plasma of number density n_0 and temperature T . The plate is located in the x, y plane and is held at a potential V_0 (negative) relative to the potential at a large distance from the plate. By following a procedure very similar to the Debye sheath problem, show that the potential in the plasma is given by

$$\Phi = V_0 \exp[-z/\lambda_D]$$

where λ_D is the Debye length.

- 2.6. Show that the number of electrons per Debye cube is given by

$$N_D = 328 \frac{T_e^{3/2}}{n_e^{1/2}},$$

where T_e is in K and n_e is in cm^{-3} .

- 2.7. Show that the boundary between a classical plasma and a quantum plasma is given by

$$T_e^{3/2} = (6.33 \times 10^{-19})n_e,$$

where T_e is in K and n_e is in cm^{-3} .

References

- Debye, P., and Hückel, E. 1923. On the Debye length in strong electrolytes. *Physikal. Z.* **24**(9), 185–206.
- Huang, K. 1963. *Statistical Mechanics*. Wiley: New York, p. 157.
- Merzbacher, E. 1998. *Quantum Mechanics*. Wiley: New York, pp. 217–220.
- Tonks, L., and Langmuir, I. 1929. Oscillations in ionized gases. *Phys. Rev.* **33**, 195–210.

Further reading

- Chen, F. F. *Introduction to Plasma Physics and Controlled Fusion*, Vol. 1: *Plasma Physics*. Plenum Press: New York, 1983 (reprinted in 1990), Chapter 1.
- Nicholson, D. R. *Introduction to Plasma Theory*. Krieger Publishing: Malabar, FL, 1992, Chapter 1 (originally published by Wiley: New York, 1983).

3

Single particle motions

A complete mathematical model of a plasma requires three basic elements: first, the motion of all particles must be determined for some assumed electric and magnetic field configuration; second, the current and charge densities must be computed from the particle trajectories; and third, the electric and magnetic fields must be self-consistently determined from the currents and charges, taking into account both internal and external sources. To be self-consistent, the electric and magnetic fields obtained from the last step must correspond to the fields used in the first step. It is this self-consistency requirement that makes the analysis of a plasma difficult.

To develop an understanding of the processes occurring in a plasma, a useful first step is to forget about the self-consistency requirement and concentrate on the motion of a single particle in a specified field configuration. This approach can be useful in a variety of situations. If the external fields are very strong and the plasma is sufficiently tenuous, the internally generated fields are sometimes small and can be safely ignored. This situation arises, for example, in radiation belts at high energies and in various electronic devices such as vacuum tubes and traveling wave amplifiers. In other situations the self-consistent electric and magnetic fields may be known from direct measurement. In this case, it is often useful to follow the motion of individual tracer particles in the known electric and magnetic fields in order to gain insight into the physical processes involved, such as particle transport and energization. Finally, in some cases it is possible to use the general solution for the particle motion in an assumed field geometry to determine a fully self-consistent solution in which the currents and charges produce the assumed fields.

3.1 Motion in a static uniform magnetic field

The simplest field configuration of importance in plasma physics is a static, uniform magnetic field. The equation of motion for a particle moving at non-relativistic velocities in a static uniform magnetic field is given by

$$m \frac{d\mathbf{v}}{dt} = q(\mathbf{v} \times \mathbf{B}), \quad (3.1.1)$$

where $q(\mathbf{v} \times \mathbf{B})$ is the Lorentz force. Taking the scalar product of the above equation with the velocity vector gives

$$m\mathbf{v} \cdot \frac{d\mathbf{v}}{dt} = \frac{d}{dt} \left(\frac{1}{2} m v^2 \right) = q\mathbf{v} \cdot (\mathbf{v} \times \mathbf{B}) = 0, \quad (3.1.2)$$

which shows that the kinetic energy, $w = (1/2)mv^2$, is a constant of the motion. To determine the particle trajectory it is convenient to resolve the velocity into components parallel, \mathbf{v}_{\parallel} , and perpendicular, \mathbf{v}_{\perp} , to the magnetic field, $\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp}$. The kinetic energy can then be written as $w = w_{\parallel} + w_{\perp}$, where $w_{\parallel} = (1/2)mv_{\parallel}^2$ and $w_{\perp} = (1/2)mv_{\perp}^2$. Since the $\mathbf{v} \times \mathbf{B}$ force has no component parallel to the magnetic field, the parallel component of the velocity is constant, so the particle moves at a constant velocity along the magnetic field. Since w and w_{\parallel} are both constants, it follows that v_{\perp} , and hence v_{\perp} , are also constants of the motion. The radius of curvature, ρ_c , of the motion in a plane perpendicular to the magnetic field can be obtained from the perpendicular component of the equation of motion (3.1.1) which, ignoring signs, can be written

$$m \frac{v_{\perp}^2}{\rho_c} = |q|v_{\perp}B. \quad (3.1.3)$$

The radius ρ_c is often called the cyclotron radius and is given by

$$\rho_c = \frac{mv_{\perp}}{|q|B}. \quad (3.1.4)$$

Since v_{\perp} is constant, the cyclotron radius is constant, which means that the motion in a plane perpendicular to the magnetic field is a circle. The frequency of the circular motion around the magnetic field, known as the cyclotron frequency, is given by

$$\omega_c = \frac{v_{\perp}}{\rho_c} = \frac{|q|B}{m}. \quad (3.1.5)$$

From the above analysis it is easy to see that the uniform motion along the magnetic field plus the circular motion around the magnetic field generates a helical trajectory, as shown in Figure 3.1. Positively charged particles rotate in the left-hand sense with respect to the magnetic field, and negatively charged particles rotate in the right-hand sense. The instantaneous center of the rotational motion is called the *guiding center*. The concept of an instantaneous guiding center will later be useful for describing motions in non-uniform magnetic fields. For a static uniform magnetic field the guiding center moves with constant velocity along the magnetic field.

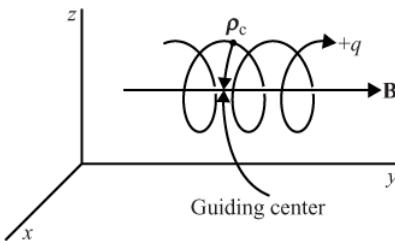


Fig. 3.1 The motion of a charged particle in a static magnetic field.

3.1.1 Magnetic moment

Since a moving charge produces a current, it is evident that the uniform circular motion around the magnetic field constitutes a current loop. The magnetic moment μ of this current loop is the product of the current and the area of the loop,

$$\mu = IA. \quad (3.1.6)$$

The magnitude of the current is simply the charge divided by the cyclotron period, and is given by

$$I = \frac{|q|}{T_c} = \frac{q^2 B}{2\pi m}. \quad (3.1.7)$$

The area of the current loop is the area of the circular orbit, which is given by

$$A = \pi \rho_c^2 = \pi \left(\frac{mv_\perp}{qB} \right)^2. \quad (3.1.8)$$

Combining equations (3.1.7) and (3.1.8), the magnetic moment can be written in the form

$$\mu = \frac{mv_\perp^2}{2B} = \frac{w_\perp}{B}. \quad (3.1.9)$$

From the direction of rotation it is also easily verified that the magnetic moment is directed opposite to the magnetic field, as indicated in Figure 3.2. This relationship holds for both positively and negatively charged particles. (Note that if q is negative the direction of rotation is reversed, but the direction of the current remains the same.) Since the magnetic moment is oriented opposite to the magnetic field, a magnetized plasma is always diamagnetic. In the presence of a large number of particles, diamagnetism appreciably reduces the magnitude of an externally imposed magnetic field. The magnetization, \mathbf{M} , defined as the magnetic moment per unit volume, is obtained by adding the magnetic moments of all species present in the plasma,

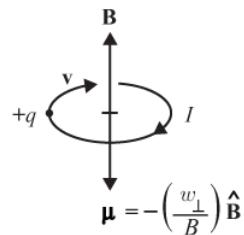


Fig. 3.2 The magnetic moment μ is directed opposite the magnetic field \mathbf{B} .

$$\mathbf{M} = - \sum_s n_s \langle \mu_s \rangle \hat{\mathbf{B}}, \quad (3.1.10)$$

where $\langle \mu_s \rangle$ is the average magnetic moment of the s th species. (Note that the average magnetic moment is directly proportional to the average perpendicular kinetic energy.) The corresponding magnetization current is given by $\mathbf{J}_m = \nabla \times \mathbf{M}$. An example illustrating this magnetization current is given in Section 3.3.4.

3.2 Motion in perpendicular electric and magnetic fields

Next we consider the more complicated case of spatially uniform, static electric and magnetic fields. For the moment, we restrict our attention to situations in which the electric field is perpendicular to the magnetic field. (If the parallel component of the electric field, E_{\parallel} , is non-zero, it will accelerate the guiding center of the particle parallel to the field line.) The equation of motion for a particle moving in such a configuration can be resolved into components parallel and perpendicular to the magnetic field, i.e.,

$$m \frac{dv_{\parallel}}{dt} = 0 \quad (3.2.1)$$

and

$$m \frac{dv_{\perp}}{dt} = q[\mathbf{E} + \mathbf{v}_{\perp} \times \mathbf{B}]. \quad (3.2.2)$$

We assume that the electric and magnetic fields are constant. By transforming to a frame of reference moving perpendicular to the magnetic field, it is possible, under certain conditions, to eliminate the electric field from equation (3.2.2) and reduce the problem to the one solved in the previous section. For non-relativistic velocities, the electric and magnetic fields transform according to the relations

$$\mathbf{E}' = \mathbf{E} + \mathbf{v}_E \times \mathbf{B}, \quad (3.2.3)$$

$$\mathbf{B}' = \mathbf{B}, \quad (3.2.4)$$

where \mathbf{v}_E is the velocity of a frame of reference (indicated by primes) moving perpendicular to \mathbf{E} and \mathbf{B} . The transformed velocities in the moving frame are $\mathbf{v}'_{\parallel} = \mathbf{v}_{\parallel}$ and $\mathbf{v}'_{\perp} = -\mathbf{v}_E + \mathbf{v}_{\perp}$. The equations of motion in the moving frame are then

$$m \frac{dv'_{\parallel}}{dt} = 0 \quad (3.2.5)$$

and

$$m \frac{dv'_{\perp}}{dt} = q[\mathbf{E}' + \mathbf{v}'_{\perp} \times \mathbf{B}]. \quad (3.2.6)$$

As long as the velocity \mathbf{v}_E is less than the speed of light, \mathbf{v}_E can be chosen such that

$$\mathbf{E}' = \mathbf{E} + \mathbf{v}_E \times \mathbf{B} = 0, \quad (3.2.7)$$

which can be inverted, by taking the cross-product with \mathbf{B} , to yield

$$\mathbf{v}_E = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (3.2.8)$$

The velocity \mathbf{v}_E is called the $\mathbf{E} \times \mathbf{B}$ drift velocity. In the moving frame of reference, the equation of motion (3.2.6) then reduces to

$$m \frac{d\mathbf{v}'}{dt} = q\mathbf{v}' \times \mathbf{B}, \quad (3.2.9)$$

which simply corresponds to a particle moving in a static uniform magnetic field. Thus, the motion of the particle in the original (unprimed) frame of reference consists of a uniform circular motion around the magnetic field plus a uniform translational motion in the $\mathbf{E} \times \mathbf{B}$ direction with velocity \mathbf{v}_E . As shown in Figure 3.3 for the special case $v_{||} = 0$, the particle trajectories are cycloids in the \mathbf{E}, \mathbf{v}_E plane, with positive charges rotating in the left-hand sense with respect to the magnetic field, and negative charges rotating in the right-hand sense. The guiding center moves at the $\mathbf{E} \times \mathbf{B}$ drift velocity. If the particle initially has a velocity component along the magnetic field, then the guiding center continues to move with this same velocity parallel to the magnetic field.

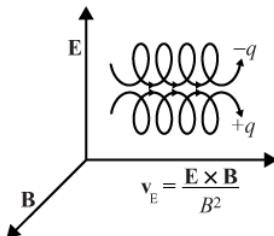


Fig. 3.3

Perpendicular electric and magnetic fields produce a drift, \mathbf{v}_E , in the $\mathbf{E} \times \mathbf{B}$ direction.

If the electric field is so large that the magnitude of \mathbf{v}_E equals or exceeds the speed of light, equation (3.2.8) is no longer valid. Under these conditions, it is not possible to transform to a reference frame in which the particle simply rotates around the magnetic field. This possibility, which rarely occurs in laboratory and space plasma applications, may be realized in some astrophysical applications (near the light cylinder at a pulsar, for example).

The motion of a charge particle in perpendicular electric and magnetic fields has several notable features.

First, all particles drift with the same velocity irrespective of their charge, mass and energy. Thus, the only effect of adding a constant perpendicular electric field is that all the particles move at the $\mathbf{E} \times \mathbf{B}$ drift velocity. This simple result occurs because the electric field is determined by the frame of reference, and in the frame moving with the $\mathbf{E} \times \mathbf{B}$ drift

velocity the electric field vanishes and the particles simply rotate about the magnetic field. Based on this result the rest frame of the plasma is usually defined as the frame of reference in which $\mathbf{E} = 0$. Note that in a plasma with an inhomogeneous electric field, this definition only applies locally, since in general it is not possible to find a frame of reference for which $\mathbf{E} = 0$ everywhere.

Second, since all particles drift at the same velocity, and since the plasma is assumed to be electrically neutral ($\sum_s n_s e_s = 0$), the $\mathbf{E} \times \mathbf{B}$ drift produces no net current density

$$\mathbf{J} = \sum_s e_s n_s \mathbf{v}_{Es} = \mathbf{v}_E \left(\sum_s n_s e_s \right) = 0. \quad (3.2.10)$$

Third, since contours of constant electrostatic potential are perpendicular to \mathbf{E} , the $\mathbf{E} \times \mathbf{B}$ drift velocity is always along a contour of constant electrostatic potential. This relationship is illustrated in Figure 3.4, which shows a magnetic field \mathbf{B} out of the paper and an inhomogeneous electric field \mathbf{E} in the plane of the paper. Contours of constant electrostatic potential, which must be perpendicular to \mathbf{E} , are indicated by Φ_1 , Φ_2 , and Φ_3 . Since the $\mathbf{E} \times \mathbf{B}$ drift is along an equipotential contour, the total energy of the particle is not changed by the $\mathbf{E} \times \mathbf{B}$ drift motion. Later, we consider other types of drifts that can cause the guiding center to move away from an equipotential contour, and thus change the kinetic energy of the particle.

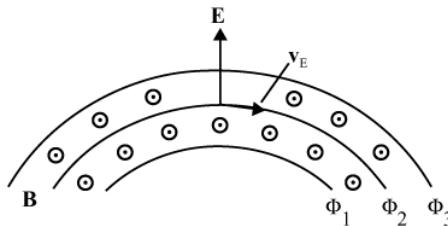


Fig. 3.4 In an inhomogeneous electric field, the $\mathbf{E} \times \mathbf{B}$ drift is along a contour of constant electrostatic potential.

3.2.1 Examples of $\mathbf{E} \times \mathbf{B}$ drifts

Numerous examples of $\mathbf{E} \times \mathbf{B}$ drifts occur in both space and laboratory plasmas. In planetary magnetospheres the collisional coupling of the plasma to the neutral atmosphere causes the plasma in the inner regions of the magnetosphere to rotate with the planet. This rotation is associated with the so-called co-rotational electric field, $\mathbf{E} = -\mathbf{v}_E \times \mathbf{B} = -(\omega \times \mathbf{r}) \times \mathbf{B}$, induced by the rotation of the planet, where ω is the angular rotation velocity of the planet. In the Earth's magnetosphere the co-rotational electric field is the dominant electric field for magnetic field lines extending out to about three to four Earth radii. The region where co-rotation dominates is called the plasmasphere (see Figure 3.5), and the outer boundary of the plasmasphere is called the plasmapause. In strongly magnetized planets, such as Jupiter, Saturn, Uranus, and Neptune, most of the magnetosphere rotates with the planet.

Co-rotational electric fields are also important in laboratory plasmas. In the laboratory, a rotating column of plasma can be produced by imposing a radial electric field on an axial magnetic field, as shown in Figure 3.6. Such a radial electric field can be produced by electrically biasing a system of concentric grids at the ends of the column. The plasma then rotates at the $\mathbf{E} \times \mathbf{B}$ velocity. If the electric field increases linearly with radial distance from the central axis, the plasma column can be made to rotate as though it were a solid body. This type of rotation is similar to magnetospheric co-rotation, except the electric field is imposed by external grids rather than by an interaction with the rotating atmosphere. In some types of magnetic confinement devices, radial electric fields arise due to charge accumulation in the plasma. External bias grids or metal plates at the ends of the plasma column can be used to control the amplitude of these fields.

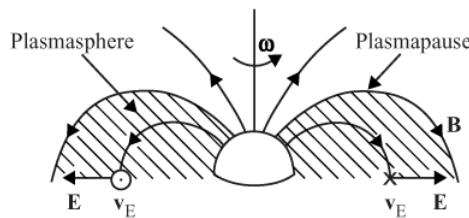


Fig. 3.5 The co-rotational electric field forces the plasma to rotate with the planet inside of a boundary called the plasmasphere.

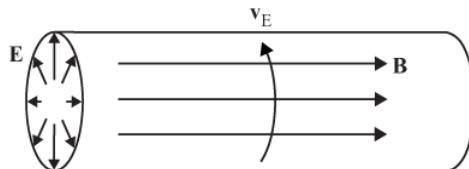


Fig. 3.6 A magnetized plasma column can be forced to rotate by imposing a radial electric field at the ends of the column.

The relationship between perpendicular electric fields and plasma drifts provides a useful method for measuring plasma flow velocities. In both laboratory and space plasmas the flow velocity perpendicular to the magnetic field can be determined by measuring the perpendicular electric field and then using $v_E = E \times B / B^2$ to compute the flow velocity. In laboratory plasmas the electric field can be determined by using a conducting probe, called a Langmuir probe, to measure the electrostatic potential at various points in the plasma relative to some reference potential. In space plasmas two probes mounted on the ends of booms are used as shown in Figure 3.7. These probes are called double probes. The electric field along the axis of the boom is computed using $E = -(\Phi_2 - \Phi_1)/L$, where Φ_1 and Φ_2 are the potentials of the two probes, and L is the separation distance. By making the probe separation large, typically 100 to 200 m, it is possible to measure very small electric fields, sometimes only a fraction of a millivolt per meter.

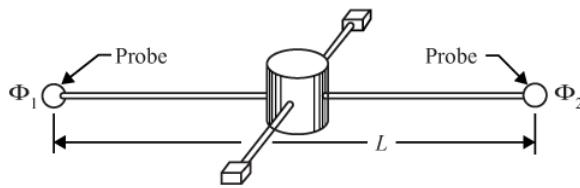


Fig. 3.7 Electric fields in space plasmas are often determined by measuring the potential difference between two probes.

3.2.2 Drift due to a force perpendicular to \mathbf{B}

The $\mathbf{E} \times \mathbf{B}$ drift can be easily generalized to include drifts due to an external force, provided the force is perpendicular to the magnetic field. If we define an equivalent electric field $\mathbf{E}_\perp = \mathbf{F}_\perp/q$ and substitute it in formula (3.2.8), we obtain the formula for the drift due to a perpendicular external force

$$\mathbf{v}_F = \frac{\mathbf{F}_\perp \times \mathbf{B}}{qB^2}. \quad (3.2.11)$$

A common example is the gravitational force $\mathbf{F} = mg$. Note that a gravitational force acting perpendicular to the magnetic field does not directly cause the plasma to “fall” because the induced drift velocity is perpendicular to the applied force. The plasma does indeed fall, but the explanation is more complicated. Unlike the $\mathbf{E} \times \mathbf{B}$ drift, the drift velocity produced by a gravitational force causes positive and negative charges to move in opposite directions, thereby causing a current. For a plasma of finite size this current produces a polarization charge at the boundaries of the plasma which, in turn, produces an electric field perpendicular to the gravitational force, as shown in Figure 3.8. This electric field causes an $\mathbf{E} \times \mathbf{B}$ drift in the direction of the gravitational force. Thus, although the drift produced by a gravitational force does not directly cause the plasma to fall, the resulting polarization electric field leads to a downward acceleration. The extent to which a polarization electric field actually develops depends on the boundary conditions. For example, if some mechanism exists to remove the polarization charge, such as grids or metal walls at the sides of the chamber, the polarization electric field can be strongly inhibited. In most laboratory and space plasmas, the effect of the gravitational force turns out to be small compared to the electric and magnetic forces.

3.3 Gradient and curvature drifts

If the magnetic field has a transverse gradient or is curved, the guiding center tends to move perpendicular to the magnetic field because of variations in the instantaneous radius of curvature. For small gradients and curvatures, the guiding center drifts with a nearly constant velocity, very similar to the $\mathbf{F} \times \mathbf{B}$ drift. The resulting drifts are called gradient

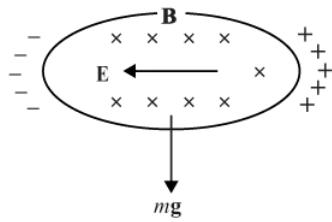


Fig. 3.8 A downward gravitational field mg causes electrons and ions to drift in opposite directions, thereby producing polarization charges at the boundaries of the plasma. The resulting electric field \mathbf{E} then causes a downward $\mathbf{E} \times \mathbf{B}$ drift.

and curvature drifts. Because these drifts have somewhat different characteristics, they are discussed separately.

3.3.1 Gradient drift

To analyze the motion of a particle moving in a magnetic field with a transverse gradient, we consider the idealized case of a magnetic field in the z direction with a magnetic field gradient, $\nabla|B|$, in the x direction. The motion of a positively charged particle in this magnetic field is sketched in Figure 3.9. As can be seen from equation (3.1.4), the instantaneous radius of curvature decreases in the region of stronger magnetic field and increases in the region of weaker magnetic field. The trajectory of the particle then consists of a cycloid-like drift motion along the y axis, perpendicular to the gradient of the magnetic field. A simple description of the orbit can be obtained if the magnetic field variation over the cyclotron orbit is small. The motion is then nearly circular in a frame of reference moving at the drift velocity.

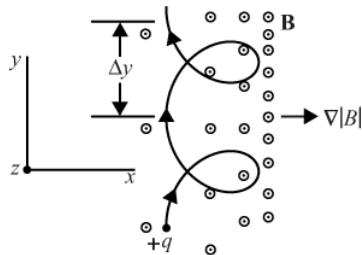


Fig. 3.9 A gradient in the magnetic field, $\nabla|B|$, causes a drift perpendicular to the gradient.

To derive the average drift velocity we make use of the fact that the motion in the x direction is periodic, which implies that the x component of the $\mathbf{v} \times \mathbf{B}$ force, time-averaged over one orbit, is zero, that is,

$$\oint F_x dt = q \oint v_y B_z dt = 0. \quad (3.3.1)$$

Since the gradient in the magnetic field strength is small, the magnetic field can be ex-

panded about the guiding center using a Taylor series. Keeping only the first-order term, the magnetic field can be written

$$B_z(x) = B_z(x_0) + \frac{\partial B_z}{\partial x}(x - x_0) + \dots \quad (3.3.2)$$

Substituting $B_z(x)$ into equation (3.3.1) and keeping only the first two terms, we obtain

$$B_z(x_0) \oint v_y dt + \frac{\partial B_z}{\partial x} \oint v_y(x - x_0) dt = 0. \quad (3.3.3)$$

The integral in the first term gives Δy , the distance moved in one orbit. The integral in the second term is the negative of the area of the near-circular orbit, so that

$$\oint v_y(x - x_0) dt = \oint (x - x_0) dy = -\frac{q}{|q|} \pi \rho_c^2. \quad (3.3.4)$$

In the above equation the $q/|q|$ term occurs because the direction of integration in the x, y plane depends on the sign of the charge. The gradient drift velocity, v_G , can then be written as

$$v_G = \frac{\Delta y}{\Delta t} = \frac{1}{\Delta t} \frac{1}{B_z} \frac{\partial B_z}{\partial x} \left(\frac{q}{|q|} \pi \rho_c^2 \right), \quad (3.3.5)$$

which, using $\Delta t = 2\pi/\omega_c$ and $\omega_\perp = (m/2)\omega_c^2\rho_c^2$, simplifies to

$$v_G = \frac{\omega_\perp}{qB_z} \left[\frac{1}{B_z} \frac{\partial B_z}{\partial x} \right]. \quad (3.3.6)$$

This expression can be rewritten in a more general vector form as

$$v_G = \frac{\omega_\perp}{qB} \left[\frac{\hat{\mathbf{B}} \times \nabla B}{B} \right], \quad (3.3.7)$$

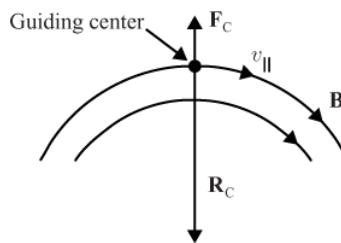


Fig. 3.10 The centrifugal force of a particle moving along a curved magnetic field causes a drift motion in the $\mathbf{B} \times \mathbf{R}_C$ direction where \mathbf{R}_C is the instantaneous radius of curvature.

where $\hat{\mathbf{B}}$ is a unit vector in the direction of the magnetic field. Note that the term in brackets has units of inverse length and is a purely geometric factor that characterizes the gradient of the magnetic field.

3.3.2 Curvature drift

The simplest way to analyze the motion of a charged particle moving in a curved magnetic field is to notice that the particle experiences a centrifugal force $F_C = mv_{\parallel}^2/R_C$, where R_C is the radius of curvature of the magnetic field. This force is directed outward from the center of curvature as shown in Figure 3.10. Since the centrifugal force is perpendicular to the magnetic field, the curvature drift velocity, \mathbf{v}_C , can be computed using the general expression (3.2.11) for the drift due to an external force which gives

$$\mathbf{v}_C = \frac{\mathbf{F} \times \mathbf{B}}{qB^2} = -m \frac{v_{\parallel}^2}{R_C^2} \frac{\mathbf{R}_C \times \mathbf{B}}{qB^2}. \quad (3.3.8)$$

The minus sign on the right-hand side of the equation arises because the centrifugal force is directed opposite to the radius of curvature vector, \mathbf{R}_C . By introducing the parallel kinetic energy, $w_{\parallel} = (1/2)mv_{\parallel}^2$, the curvature drift can be rewritten in a form that is very similar to the equation for gradient drift

$$\mathbf{v}_C = \frac{2w_{\parallel}}{qB} \left[\frac{\hat{\mathbf{B}} \times \hat{\mathbf{R}}_C}{R_C} \right], \quad (3.3.9)$$

where $\hat{\mathbf{R}}_C$ is a unit vector in the direction of the instantaneous radius of curvature. Note that the term in the brackets has units of inverse length, and is a purely geometric factor that characterizes the curvature of the magnetic field.

Gradient and curvature drifts have many similar features. In both cases the drift velocity is proportional to an inverse length scale that characterizes the inhomogeneity of the magnetic field. For many common field geometries, such as toroidal and dipole fields, these length scales are quite similar, which means that the gradient and curvature drifts tend to have similar magnitudes. In both cases the drift velocities are proportional to the kinetic energy (perpendicular energy for the gradient drift, and parallel energy for the curvature drift) and inversely proportional to the magnetic field strength. The crucial dimensionless quantity controlling the magnitude of the drift velocity is the ratio of the cyclotron radius to the characteristic length scale of the magnetic field inhomogeneity. Increasing the kinetic energy or decreasing the magnetic field strength increases the cyclotron radius and, consequently, the drift velocity.

Gradient and curvature drifts differ from the $\mathbf{E} \times \mathbf{B}$ drift in several important respects. First, the gradient and curvature drift velocities are both proportional to the kinetic energy, whereas the $\mathbf{E} \times \mathbf{B}$ drift is independent of the kinetic energy. Thus, the $\mathbf{E} \times \mathbf{B}$ drift tends to control the motion of low-energy (“cold”) particles, whereas gradient and curvature drifts tend to control the motion of high-energy (“hot”) particles. Second, the gradient and curvature drift velocities both depend on the sign of the charge, whereas the $\mathbf{E} \times \mathbf{B}$ drift is independent of the charge. Gradient and curvature drifts, therefore, give rise to currents, whereas the $\mathbf{E} \times \mathbf{B}$ drift does not. Third, gradient and curvature drifts do not necessarily occur along contours of constant electrostatic potential. Therefore, these drifts can cause particles to gain or lose energy as the particles drift onto different electrostatic potential

contours. The gain or loss of energy is due to the work done by the electric field. In general, only the “cold” plasma particles move with no change in kinetic energy.

3.3.3 Examples of gradient and curvature drifts

In the Earth’s radiation belt, energetic charged particles experience both gradient and curvature drifts. From the geometry of the dipole magnetic field, it is easy to see that the gradient and curvature drifts are in the same direction, causing electrons to drift eastward around the Earth, and causing positively charged ions to drift westward. As shown in Figure 3.11, these drifts produce a current \mathbf{J} that encircles the Earth in the westward direction. This current is called the ring current. The magnitude of the ring current depends on the number density and kinetic energy of the particles in the radiation belt. During disturbed periods, called magnetic storms, a dense energetic plasma is injected into the magnetic field, intensifying the ring current and producing measurable magnetic disturbances at the Earth. Because the ring current is proportional to the kinetic energy of the particles, the magnetic disturbance at the Earth gives a rough measure of the total energy of the particles in the ring current.

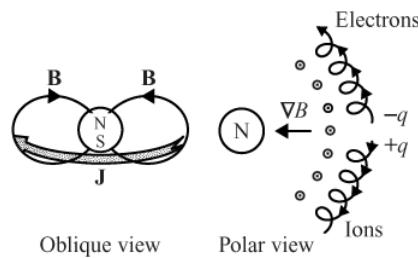


Fig. 3.11

In the Earth’s dipole magnetic field electrons drift eastward and ions drift westward, producing a current \mathbf{J} called the ring current.

As an example of gradient and curvature drifts in laboratory plasmas, consider the toroidal magnetic field. This magnetic field configuration can be produced by wrapping a current-carrying winding around the torus, as shown in Figure 3.12. Since the magnetic field lines close on themselves, it appears at first glance that the guiding center motion would consist of an azimuthal circulation around the torus on the same closed magnetic field line, thereby ensuring a long confinement time. However, a radial gradient exists in the magnetic field strength, as can be seen from Ampère’s law,

$$B = \frac{\mu_0 N_\ell I}{2\pi \rho}, \quad (3.3.10)$$

where μ_0 is the permeability of free space, N_ℓ is the number of turns per unit length, I is the current, and ρ is the distance from the symmetry axis. The gradient is directed inward, toward the center of the torus, and produces a gradient drift along the vertical symmetry axis of the torus. Since the magnetic field lines are curved, a particle spiraling along the magnetic field also experiences curvature drift. It is easily verified that the gradient and

curvature drifts are both of similar magnitude and in the same direction. Electrons and ions drift in opposite directions, as illustrated in Figure 3.13. These gradient and curvature drifts would eventually cause the particles to hit the wall of the machine. However, the actual dynamics is more complicated. Since the electrons and ions drift in opposite directions, a polarization charge soon builds up on the upper and lower boundaries of the torus. This polarization charge produces an electric field along the symmetry axis of the torus. This electric field in turn causes an $\mathbf{E} \times \mathbf{B}$ drift that is directed radially outward, causing the plasma to move radially outward until it hits the walls of the machine, where it is lost.

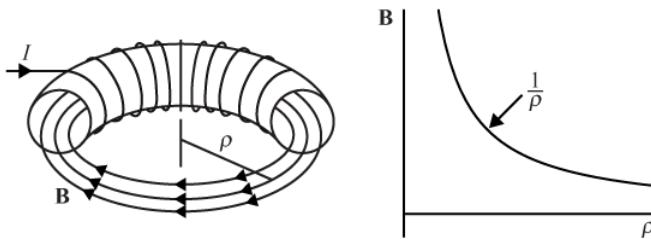


Fig. 3.12 A toroidal magnetic field has a gradient directed radially inward, toward the symmetry axis.

An ingenious solution to this particle confinement problem was found by researchers in the field of controlled thermonuclear fusion. They suggested that the toroidal magnetic field be given a twist as illustrated in Figure 3.14, either by using an external, helical current-carrying winding, as in a machine called a stellarator, or by passing an axial current through the plasma, as in a machine called a tokamak. Since polarization charges can move along the magnetic field lines from the bottom to the top of the torus, the twisted magnetic field geometry prevents the development of a polarization electric field. The $\mathbf{E} \times \mathbf{B}$ drift is then strongly inhibited, resulting in a much longer confinement time.

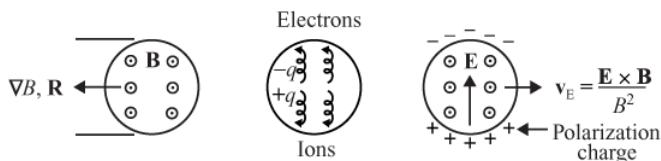


Fig. 3.13 Gradient and curvature drifts in a toroidal magnetic field cause electrons and ions to drift in opposite directions, causing polarization charges. The resulting electric field causes an outward $\mathbf{E} \times \mathbf{B}$ drift.

3.3.4 A self-consistent static equilibrium

As discussed in the beginning of this chapter, one of our objectives is to obtain a self-consistent plasma configuration. Therefore, it is instructive to analyze a simple self-consistent plasma-magnetic field configuration that involves some of the drift motions described above. To illustrate the procedure, consider an idealized two-dimensional system in which

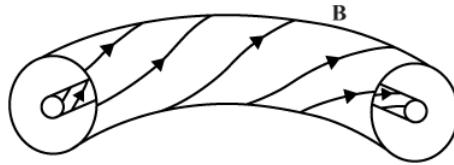


Fig. 3.14 | Polarization charges in a toroidal plasma can be inhibited by twisting the magnetic field lines.

the spatially varying magnetic field is parallel to the z axis, with $\mathbf{B} = [0, 0, B_z(x, y)]$. If a plasma is placed in such a magnetic field, two types of currents occur, a magnetization current, \mathbf{J}_m , and a gradient drift current, \mathbf{J}_G . The magnetization current can be computed from the magnetization, \mathbf{M} , which is given by

$$\mathbf{M} = - \sum_s n_s \langle \mu_s \rangle \hat{\mathbf{B}} = -\hat{\mathbf{z}} \sum_s n_s \frac{\langle w_{\perp s} \rangle}{B} = -\hat{\mathbf{z}} \frac{W_{\perp}}{B}, \quad (3.3.11)$$

where $W_{\perp} = \sum_s n_s \langle w_{\perp s} \rangle$ is the average perpendicular kinetic energy density, summed over all species. The magnetization current is given by

$$\mathbf{J}_m = \nabla \times \mathbf{M} = \hat{\mathbf{z}} \times \nabla \left(\frac{W_{\perp}}{B} \right) = \frac{\hat{\mathbf{z}}}{B} \times \nabla W_{\perp} - \frac{W_{\perp}}{B^2} \hat{\mathbf{z}} \times \nabla B, \quad (3.3.12)$$

where we have used the vector identity $\nabla \times (\phi \mathbf{F}) = \nabla \phi \times \mathbf{F} + \phi \nabla \times \mathbf{F}$. The gradient drift current is given by

$$\mathbf{J}_G = \sum_s n_s e_s \langle \mathbf{v}_{Gs} \rangle = \sum_s n_s e_s \frac{\langle w_{\perp s} \rangle}{e_s B^2} (\hat{\mathbf{z}} \times \nabla B), \quad (3.3.13)$$

which can be simplified to

$$\mathbf{J}_G = \frac{W_{\perp}}{B^2} \hat{\mathbf{z}} \times \nabla B. \quad (3.3.14)$$

The total current is then

$$\mathbf{J} = \mathbf{J}_m + \mathbf{J}_G = \frac{\hat{\mathbf{z}}}{B} \times \nabla W_{\perp}. \quad (3.3.15)$$

To provide a self-consistent, static equilibrium the above current must be related to the magnetic field via Ampère's law, so that

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} = \frac{\mu_0}{B} \hat{\mathbf{z}} \times \nabla W_{\perp}. \quad (3.3.16)$$

Using the vector identity $\nabla \times (\hat{\mathbf{z}} B) = -\hat{\mathbf{z}} \times \nabla B$, this equation can be written

$$\hat{\mathbf{z}} \times \left(\nabla B + \frac{\mu_0}{B} \nabla W_{\perp} \right) = 0. \quad (3.3.17)$$

Noting that $\nabla B^2 = 2B\nabla B$, the above equation can be simplified to

$$\hat{\mathbf{z}} \times \nabla \left(\frac{B^2}{2\mu_0} + W_{\perp} \right) = 0, \quad (3.3.18)$$

which is satisfied if

$$\frac{B^2}{2\mu_0} + W_{\perp} = \text{constant}. \quad (3.3.19)$$

It is readily verified that both $B^2/2\mu_0$ and W_{\perp} have units of pressure. Later we will demonstrate that condition (3.3.19) is a special case of a more general magnetostatic equilibrium condition, in which the sum of the magnetic field pressure, $B^2/2\mu_0$, and the particle pressure, W_{\perp} , is a constant.

3.4 Motion in a magnetic mirror field

A magnetic mirror field is characterized by a spatial variation in the magnetic field strength along the magnetic field lines. As will be shown, this type of spatial variation leads to a reflection in the parallel component of the guiding center motion, which is why the field is called a mirror field. An example of an axially symmetric, magnetic mirror field is shown in Figure 3.15. For an axially symmetric field, it is convenient to use cylindrical coordinates, with ρ representing the radius from the symmetry axis, ϕ representing the azimuthal angle around the symmetry axis, and z representing the position along the symmetry axis.

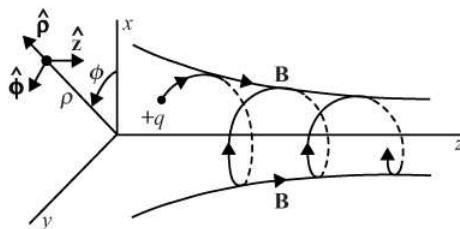


Fig. 3.15 The coordinate system used for analyzing the motion of a charged particle in a magnetic mirror field.

If the magnetic field strength increases along the magnetic field line, then Maxwell's equation $\nabla \cdot \mathbf{B} = 0$ shows that the field lines must converge as one moves along the axis of symmetry. The rate of convergence can be determined from Maxwell's equation

$$\nabla \cdot \mathbf{B} = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho B_{\rho} \right) + \frac{\partial B_z}{\partial z} = 0, \quad (3.4.1)$$

which after integrating once, assuming that $\partial B_z / \partial z$ is independent of ρ , gives

$$B\rho = -\frac{1}{2} \left(\frac{\partial B_z}{\partial z} \right) \rho \quad (3.4.2)$$

or

$$B_x = -\frac{1}{2} \left(\frac{\partial B_z}{\partial z} \right) x \text{ and } B_y = -\frac{1}{2} \left(\frac{\partial B_z}{\partial z} \right) y. \quad (3.4.3)$$

If $\partial B_z / \partial z$ is positive, the field lines converge towards the symmetry axis as z increases. These equations are valid only near the axis of symmetry, since in doing the integration we have assumed that $\partial B_z / \partial z$ is independent of ρ .

When analyzing the Lorentz force on a particle moving in a magnetic mirror field, it is convenient to consider the parallel and azimuthal motions separately.

3.4.1 Parallel motion: the magnetic mirror force

It is easy to show that the equation for the parallel motion along the z axis is given by

$$m \frac{dv_z}{dt} = F_z = q [v_x B_y - v_y B_x]. \quad (3.4.4)$$

Using equation (3.4.3), the above equation can be written

$$F_z = -\frac{q}{2} \left(\frac{\partial B_z}{\partial z} \right) (v_x y - v_y x). \quad (3.4.5)$$

If the magnetic field varies sufficiently slowly, the transverse motion is very nearly circular, in which case, to a good approximation, we can write

$$x = \rho_c \sin : \omega_c t, \quad y = \frac{q}{|q|} \rho_c \cos \omega_c t \quad (3.4.6)$$

and

$$v_x = \omega_c \rho_c \cos \omega_c t, \quad v_y = -\frac{q}{|q|} \omega_c \rho_c \sin \omega_c t. \quad (3.4.7)$$

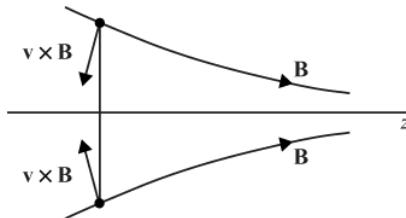


Fig. 3.16 In a converging magnetic field the $v \times B$ force has a component directed away from the region of stronger magnetic field.

The z component of the force then becomes

$$F_z = -\frac{\partial B_z}{\partial z} \left(\frac{|q|}{2} \omega_c \rho_c^2 \right). \quad (3.4.8)$$

It is easy to show that the term inside the parentheses, $|q|\omega_c\rho_c^2/2$, is simply the magnetic moment $\mu = w_\perp/B$. Thus, we obtain the following simple equation for the z component of the force

$$F_z = -\mu \frac{\partial B_z}{\partial z}. \quad (3.4.9)$$

If we let $\mathbf{m} = -\mu \hat{\mathbf{B}}$, this equation reduces to the well-known equation for the force on a magnetic moment in an inhomogeneous magnetic field, $\mathbf{F} = (\mathbf{m} \cdot \nabla) \mathbf{B}$. Thus, the parallel force arises from the interaction between the magnetic field and the magnetic moment produced by the cyclotron motion of the particle. The direction of the force is always such that the particle tends to be repelled from the region of strong magnetic field. The physical origin of this force is a consequence of the convergence of the magnetic field lines. As shown in Figure 3.16, if the magnetic field lines are converging, the Lorentz ($\mathbf{v} \times \mathbf{B}$) force has a component opposite to the direction of convergence, thereby producing a force away from the region of stronger magnetic field.

3.4.2 Azimuthal motion: constancy of the magnetic moment

In addition to the parallel component, the Lorentz force also has a component in the azimuthal direction. This force is given by

$$F_\phi = q v_z B_\rho, \quad (3.4.10)$$

and is directed as shown in Figure 3.17. The azimuthal force produces a torque that causes the perpendicular kinetic energy to change as the particle moves along the z axis. The rate of change of the perpendicular kinetic energy can be computed from the rate at which work is done by the azimuthal force,

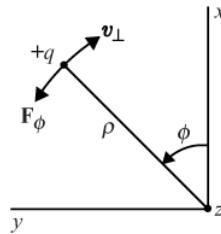


Fig. 3.17 The coordinate system used to evaluate the azimuthal torque. Note that for a positive charge v_\perp is in the minus $\dot{\phi}$ direction, so $v_\phi = -v_\perp$.

the z axis. The rate of change of the perpendicular kinetic energy can be computed from the rate at which work is done by the azimuthal force,

$$\frac{dw_{\perp}}{dt} = v_{\phi}F_{\phi} = qv_{\phi}v_zB_{\rho}. \quad (3.4.11)$$

Substituting equation (3.4.2) for B_{ρ} and noting that $v_{\phi} = -(q/|q|)v_{\perp}$, the above equation becomes

$$\frac{dw_{\perp}}{dt} = |q|v_{\perp}v_z\left(\frac{\partial B_z}{\partial z}\right)\frac{\rho}{2}. \quad (3.4.12)$$

Note that although the perpendicular kinetic energy changes, the total kinetic energy remains constant because the $\mathbf{v} \times \mathbf{B}$ force is always perpendicular to \mathbf{v} . Using the cyclotron radius $\rho_c = mv_{\perp}/|q|B$ for ρ , the rate of change of the perpendicular kinetic energy can be written

$$\frac{dw_{\perp}}{dt} = \frac{w_{\perp}v_z}{B}\left(\frac{\partial B_z}{\partial z}\right). \quad (3.4.13)$$

From the above result, we next show that the magnetic moment is a constant of the motion. The rate of change of the magnetic moment is given by

$$\frac{d\mu}{dt} = \frac{d}{dt}\left(\frac{w_{\perp}}{B}\right) = \frac{1}{B}\frac{dw_{\perp}}{dt} - \frac{w_{\perp}}{B^2}\frac{dB}{dt}, \quad (3.4.14)$$

which, since $dB/dt = v_z(\partial B/\partial z)$, can be written

$$\frac{d\mu}{dt} = \frac{1}{B}\frac{dw_{\perp}}{dt} - \frac{w_{\perp}v_z}{B^2}\left(\frac{\partial B_z}{\partial z}\right). \quad (3.4.15)$$

Substituting expression (3.4.13) for dw_{\perp}/dt into the above equation, we obtain or $\mu = \text{constant}$. As in the previous section, this result is valid only if the magnetic field varies slowly, which is why the magnetic moment is called an adiabatic invariant. Later, a more specific criterion will be developed to determine how slow this variation must be.

$$\frac{d\mu}{dt} = \frac{w_{\perp}v_z}{B^2}\left(\frac{\partial B_z}{\partial z}\right) - \frac{w_{\perp}v_z}{B^2}\left(\frac{\partial B_z}{\partial z}\right) = 0, \quad (3.4.16)$$

3.4.3 Turning points and the pitch angle

The parallel equation of motion ((3.4.4) together with (3.4.9)), which was derived for a small local region, can be easily extended to an arbitrary magnetic field configuration by simply changing z to s , the distance along the field line, v_z to $v_{||}$, the parallel component of velocity, and B_z to B , the magnitude of the magnetic field. With these substitutions, the parallel equation of motion becomes

$$m\frac{dv_{||}}{dt} = -\mu\frac{\partial B}{\partial s}, \quad (3.4.17)$$

where B is the magnetic field strength at position s . Writing $dv_{\parallel}/dt = v_{\parallel} dv_{\parallel}/ds$ and noting that μ is constant, the above equation can be integrated once to give

$$\frac{d}{ds} \left(\frac{1}{2}mv_{\parallel}^2 + \mu B \right) = 0, \quad (3.4.18)$$

which implies that

$$\frac{1}{2}mv_{\parallel}^2 + \mu B(s) = \mu B_m, \quad (3.4.19)$$

where μ and B_m are constants. The above equation can be interpreted as an energy conservation equation for a particle of mass m moving in a one-dimensional effective potential $\mu B(s)$. The effective potential for a typical magnetic mirror field is shown in Figure 3.18. The turning points of the motion are at s_1^* and s_2^* , where $B = B_m$. Such a configuration is often called a magnetic bottle, since the particle is trapped between the two turning points. Since the parallel kinetic energy at the turning points is zero, it is easy to see that μB_m is the total kinetic energy of the particle, $w = \mu B_m$. The energy conservation equation can then be solved for the parallel velocity

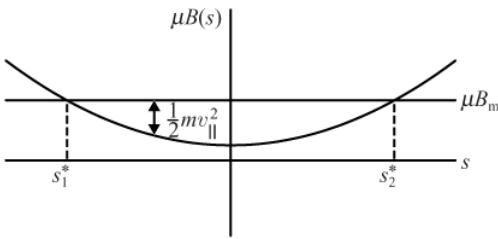


Fig. 3.18 The effective potential for motion along the magnetic field.

$$v_{\parallel} = \frac{ds}{dt} = \pm \sqrt{\frac{2\mu}{m} (B_m - B)}, \quad (3.4.20)$$

which in turn can be integrated to determine completely the motion of the particle along the field line.

It is often useful to know the angle between the velocity vector and the magnetic field. This angle is called the pitch angle α . The pitch angle at any point can be computed from the magnetic moment by noting that $v_{\perp} = v \sin \alpha$, which gives

$$\mu = \frac{w_{\perp}}{B} = \frac{w \sin^2 \alpha}{B}. \quad (3.4.21)$$

Since $w = \mu B_m$ and $w_{\perp} = \mu B$, the above equation simplifies to

$$\sin^2 \alpha = \frac{B}{B_m}. \quad (3.4.22)$$

For a particle of given energy and magnetic moment, this equation can be used to compute the pitch angle at any point along the magnetic field line.

Practical considerations dictate that in any magnetic mirror configuration there is always an upper limit to the magnetic field strength at the ends or “throats” of the magnetic bottle. This upper limit implies that a minimum pitch angle exists below which particles are not reflected, but are lost from the bottle. If the minimum field strength along a field line at the mid-plane is defined to be B_0 as shown in Figure 3.19, it is easy to show that the minimum pitch angle α_0 , below which particles are lost, is given by

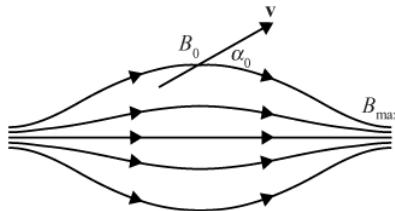


Fig. 3.19 The magnetic field in an axially symmetric mirror machine always has an upper limit to the magnetic field, B_{\max} , at the ends of the mirror. This upper limit causes a lower limit to the pitch angle, α_0 , at the middle of the machine. Particles with pitch angles less than α_0 are lost from the machine.

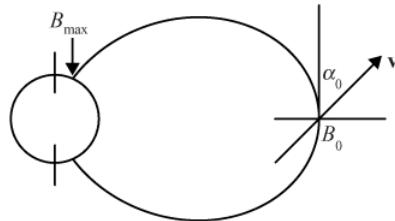


Fig. 3.20 Planetary radiation belts also have a lower limit to the pitch angle, α_0 . Particles with pitch angles less than α_0 strike the planet and are lost from the radiation belt.

$$\sin^2 \alpha_0 = \frac{B_0}{B_{\max}}, \quad (3.4.23)$$

where B_{\max} is the maximum magnetic field strength. Any particle with a pitch angle less than α_0 will escape from the system. The cone of directions around the magnetic field at the angle α_0 is called the loss cone. Later we will show that a loss cone in the particle velocity distribution leads to instabilities that have detrimental effects on the confinement of a plasma in a magnetic mirror machine. Charged particles trapped in planetary radiation belts also have a loss cone because any particle moving with a pitch angle less than some minimum angle α_0 will hit the planet and be lost. The geometry involved is illustrated in Figure 3.20. It is easy to see that α_0 is also given by equation (3.4.23) where B_{\max} is now the magnetic field strength at the surface of the planet.

3.5 Motion in a time varying magnetic field

If the magnetic field varies with time, Faraday's law, $\nabla \times \mathbf{E} = -\partial \mathbf{B}/\partial t$, shows that a non-conservative electric field is present. Therefore, the total energy of the particle is not constant. To illustrate the effect of the non-conservative electric field, consider the idealized case of a uniform magnetic field that is varying on a time scale which is slow compared with the cyclotron period. If the magnetic field is directed upward, out of the paper, and $\partial \mathbf{B}/\partial t > 0$, the non-conservative electric field \mathbf{E} is in the clockwise sense, as shown in Figure 3.21. A charged particle orbiting in this field gains an amount of kinetic energy per orbit given by

$$\Delta w_{\perp} = q \int_C \mathbf{E} \cdot d\ell = -q \int_S \frac{\partial \mathbf{E}}{\partial t} \cdot d\mathbf{A} \quad (3.5.1)$$

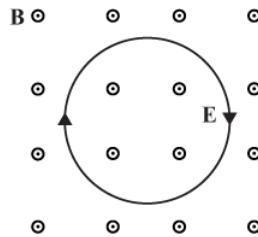


Fig. 3.21 The non-conservative electric field induced by a time varying magnetic field.

where Stokes' theorem has been used to change the line integral to an area integral and the positive sense of the curve C is in the direction of the particle motion. To a good approximation the area of the orbit is $\pi \rho_c^2$. If the magnetic field varies slowly with time, the change in the perpendicular energy during one orbit is

$$\Delta w_{\perp} = |q| \left(\frac{dB}{dt} \right) (\pi \rho_c^2), \quad (3.5.2)$$

where $|q|$ is used because the energy gain is independent of the sign of the charge. Note that for positive charges the curve C in equation (3.5.1) must be clockwise, which means that $d\mathbf{A}$ must be directed into the paper. For negative charges the direction of $d\mathbf{A}$ is reversed, which accounts for the $|q|$ in equation (3.5.2). Since the cyclotron period is given by $\Delta t = 2\pi/\omega_c$, the rate of change of the perpendicular energy is given by

$$\frac{dw_{\perp}}{dt} = \frac{1}{2\pi} \omega_c |q| \left(\frac{dB}{dt} \right) (\pi \rho_c^2), \quad (3.5.3)$$

which, noting that $\mu = (1/2\pi)\omega_c |q| \pi \rho_c^2$, can be written

$$\frac{dw_{\perp}}{dt} = \mu \left(\frac{dB}{dt} \right). \quad (3.5.4)$$

Since $\mu = w_{\perp}/B$, the equation for the perpendicular energy then becomes

$$\frac{dw_{\perp}}{w_{\perp}} = \frac{dB}{B}, \quad (3.5.5)$$

which can be integrated directly to give

$$\frac{w_{\perp}}{B} = \text{constant}. \quad (3.5.6)$$

But this ratio is just the magnetic moment, μ . Thus, for a time dependent magnetic field we again arrive at the conclusion that the magnetic moment is a constant.

The fact that the magnetic moment is constant in both time dependent and converging magnetic fields is no accident. By simply changing to a coordinate system moving with the guiding center of the particle, the motion in a converging magnetic field can be converted to a time dependent magnetic field. Thus, the azimuthal torque encountered in a magnetic mirror field can be thought of as arising from the non-conservative electric field produced by the time varying magnetic field encountered as the guiding center of the particle moves along the magnetic field line.

It is also interesting to note that conservation of the magnetic moment is equivalent to maintaining a constant magnetic flux through the cyclotron orbit. This interpretation follows from the expression for the magnetic moment which can be written

$$\mu = \frac{|q|}{2} \omega_c \rho_c^2 = \frac{1}{2\pi} \frac{q^2}{m} \Phi_B, \quad (3.5.7)$$

where

$$\Phi_B = \pi \rho_c^2 B \quad (3.5.8)$$

is the magnetic flux Φ_B through the orbit. Since $\mu = \text{constant}$, the magnetic flux through the orbit remains constant, irrespective of whether the change is due to an explicit time variation or to the motion in an inhomogeneous magnetic field. This relationship is illustrated in Figure 3.22 where $\Phi_{B1} = \Phi_{B2}$ at two different times, t_1 and t_2 , in both a time varying and a spatially varying magnetic field.

3.6 Adiabatic invariants

From Hamiltonian mechanics, it is known that in a nearly periodic system with slowly varying parameters, the action integral

$$J = \oint p \, dq, \quad (3.6.1)$$

is an adiabatic invariant [Goldstein, 1959]. Here p and q are the conjugate momentum and position coordinates, and the integral is performed around a complete cycle while holding

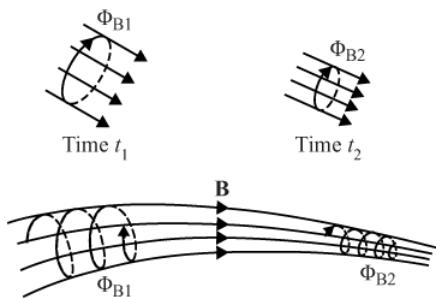


Fig. 3.22 The change in the cyclotron orbit due to a time varying magnetic field (top), and a spatially varying magnetic field (bottom). In both cases the magnetic flux through the orbit is conserved, i.e., $\Phi_{B1} = \Phi_{B2}$.

the parameters of the system constant. Adiabatic invariants are not exact invariants, but are approximately constant whenever the parameters of the system vary sufficiently slowly.

3.6.1 Adiabatic invariant for a harmonic oscillator

As a simple example it is useful to discuss the action integral for the time dependent harmonic oscillator, described by the equation

$$\frac{d^2x}{dt^2} + \omega(t)^2 x = 0. \quad (3.6.2)$$

In a spring-mass system, for instance, x represents the displacement of the mass from equilibrium when the spring is unstretched. The frequency $\omega(t)$ is determined by two system parameters, the spring constant and the particle mass. If these parameters are held fixed, the frequency $\omega(t)$ is constant, and the position $q = x$ and the conjugate momentum $p = mv_x$ are easily shown to be

$$q = A \sin \omega t \quad (3.6.3)$$

and

$$p = m\omega A \cos \omega t, \quad (3.6.4)$$

where A is the amplitude of the oscillation. The resulting trajectory in the p, q plane is an ellipse, as shown in Figure 3.23. The action integral is then simply the area of the ellipse, πab , where a and b are the minor and major radii of the ellipse, so that

$$J = \pi m \omega A^2. \quad (3.6.5)$$

If the oscillator frequency varies sufficiently slowly, the action integral J is an approximate constant of the motion (i.e., an adiabatic invariant). The amplitude of the oscillation must

then vary in such a way that $\omega A^2 \approx \text{constant}$, or $A \propto 1/\sqrt{\omega}$. Since the total energy is given by $W = (1/2)m\omega^2A^2$, the adiabatic invariance of J

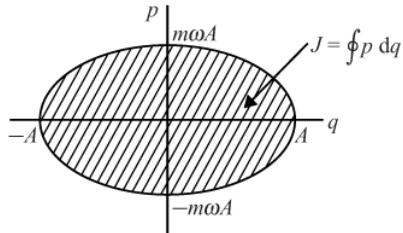


Fig. 3.23 For a harmonic oscillator, the motion in the p, q plane is an ellipse. The action integral, J , is the area of the ellipse.

is equivalent to the statement that

$$\frac{W}{\omega} = \text{constant}. \quad (3.6.6)$$

This result applies, for example, to a simple pendulum the length of which is slowly varied. If both W and ω change sufficiently slowly, the ratio of these two quantities remains essentially constant.

Although the action integral is known to be an approximate constant of the motion if the parameters of the system vary sufficiently slowly, it is interesting to ask just how slow these variations must be for the concept to be valid. This question can be answered by investigating the so-called WKB solution of the time dependent harmonic oscillator equation, which is given by

$$X_{\text{WKB}} = \frac{1}{\sqrt{\omega(t)}} \exp \left[\pm i \int^t \omega(t') dt' \right], \quad (3.6.7)$$

after Wentzel [1926], Kramers [1926], and Brillouin [1926]. Note that the amplitude factor in the WKB solution, $A = 1/\sqrt{\omega(t)}$, is such that the action integral $J = \pi m\omega A^2$ is precisely constant. It is relatively straightforward to show that X_{WKB} is an exact solution to the differential equation,

$$\frac{d^2 X_{\text{WKB}}}{dt^2} + \left[\omega^2 + \frac{\ddot{\omega}}{2\omega} - \frac{3}{4} \left(\frac{\dot{\omega}}{\omega} \right)^2 \right] X_{\text{WKB}} = 0, \quad (3.6.8)$$

where \cdot denotes the time derivative. From the above equation one can see that the WKB solution is a good solution to the time-dependent harmonic oscillator equation whenever

$$\omega^2 \gg \left| \frac{3}{4} \left(\frac{\dot{\omega}}{\omega} \right)^2 - \frac{\ddot{\omega}}{2\omega} \right|. \quad (3.6.9)$$

As a rough rule we can say that the action integral is a “good” constant of the motion if the fractional change in the oscillator frequency is small during one cycle. Note that the WKB solution always fails whenever the frequency goes through zero ($\omega = 0$).

3.6.2 Adiabatic invariants for an axially symmetric magnetic mirror

It is evident from Section 3.6.1 that an action integral exists for each degree of freedom that exhibits a periodicity. For the case of a charged particle in an axially symmetric magnetic mirror field, three types of periodic motion can be identified: (1) the cyclotron motion of the particle around the magnetic field, (2) the parallel bounce motion along the magnetic field line, and (3) the azimuthal gradient and curvature drift around the symmetry axis. These three types of motion are illustrated in Figure 3.24. The action integrals corresponding to these three motions are called the 1st, 2nd, and 3rd adiabatic invariants, respectively.

3.6.2.1 The 1st adiabatic invariant

The 1st adiabatic invariant can be obtained from the equations of motion for a particle in a uniform magnetic field B_z , which are

$$m \frac{dv_x}{dt} = qv_y B_z \text{ and } m \frac{dv_y}{dt} = -qv_x B_z. \quad (3.6.10)$$

By eliminating either v_x or v_y between these two equations, it is easy to show that the equations for the x and y motions reduce to two identical harmonic oscillator equations

$$\frac{d^2x}{dt^2} + \omega_c^2 x = 0 \text{ and } \frac{d^2y}{dt^2} + \omega_c^2 y = 0. \quad (3.6.11)$$

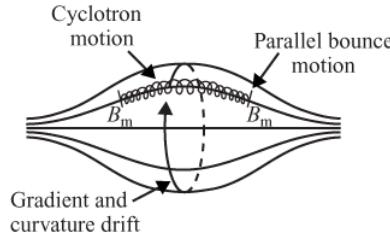


Fig. 3.24 The three types of motion for a particle trapped in an axially symmetric magnetic mirror field.

Since the amplitude of the x and y motions is just the cyclotron radius ρ_c the corresponding action integral is

$$J = \pi m \omega_c \rho_c^2. \quad (3.6.12)$$

Except for a multiplicative constant, this action integral is the same as the magnetic moment

$$\mu = \frac{|q|}{2} \omega_c \rho_c^2. \quad (3.6.13)$$

Therefore, the 1st adiabatic invariant is simply the magnetic moment, which has been shown in Sections 3.4 and 3.5 to be an approximate constant of the motion whenever the magnetic field varies sufficiently slowly.

3.6.2.2 The 2nd adiabatic invariant

The action integral for the parallel motion can be written

$$J = m \oint v_{\parallel} ds, \quad (3.6.14)$$

where the integration is along the magnetic field line between the two mirror points. The parallel velocity can be determined from the energy conservation equation (3.4.19)

$$\frac{1}{2}mv_{\parallel}^2 + \mu B(s) = \mu B_m, \quad (3.6.15)$$

so that

$$J = \sqrt{2\mu m} \oint \sqrt{B_m - B(s)} ds. \quad (3.6.16)$$

The quantity J is called the 2nd adiabatic invariant. Since μ is a constant, the 2nd adiabatic invariant is equivalent to requiring that the integral

$$\int_a^b \sqrt{B_m - B(s)} ds \quad (3.6.17)$$

be a constant, where a and b are the two mirror points.

To illustrate the use of J , consider the situation where the magnetic field is slowly changed from some initial configuration $B_i(s)$ to some final configuration $B_f(s)$. Since the magnetic field changes with time, the total energy, $w = \mu B_m$, is no longer a constant of the motion. The locations of the new mirror points can be determined from the invariance of J which requires that

$$\int_{a_i}^{b_i} \sqrt{B_{mi} - B_i(s)} ds = \int_{a_f}^{b_f} \sqrt{B_{mf} - B_f(s)} ds. \quad (3.6.18)$$

Since the initial configuration is known, the integral on the left has a known value. The final mirror point field B_{mf} and the corresponding mirror points a_f and b_f are then adjusted until the integral on the right matches the known value for the integral on the left. This procedure is illustrated schematically in Figure 3.25, where $B_i(s)$ is the initial magnetic field and $B_f(s)$ is the final magnetic field. Once the final mirror point field is determined, the final total energy is given by $w_f = \mu B_{mf}$.

The 2nd adiabatic invariant also applies to geometries where the magnetic field is not azimuthally symmetric. Since gradient and curvature drifts cause an azimuthal drift motion, the problem is usually to determine the field line that the particle is on after it has drifted to a new azimuthal location. If the magnetic field is constant in time then the total energy

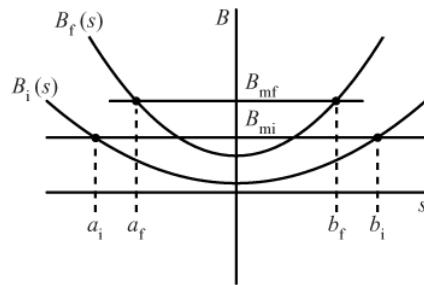


Fig. 3.25 A diagram illustrating the use of the 2nd adiabatic invariant in a slowly varying magnetic mirror field.

is conserved and the mirror point field B_m is constant. The 2nd adiabatic invariant can then be written in the form

$$I = \int_a^b \left(1 - \frac{B(s)}{B_m}\right)^{1/2} ds = \text{constant}, \quad (3.6.19)$$

where the integral I is called the integral invariant. As the particle drifts in azimuth, the locations of the new mirror points are determined by two conditions: first, that the magnetic field strength at the mirror points remain constant; and second, that the integral I remain constant. This procedure is often used to determine the drift motion of a particle trapped in a planetary magnetic field. Higher-order multipole moments typically cause planetary magnetic fields to have substantial azimuthal distortions relative to an ideal dipole field. Conservation of the 1st adiabatic invariant assures us that the magnetic field at the mirror points remains constant ($B_m = B_{mi}$). As shown in Figure 3.26, the condition $B_m = B_{mi}$ corresponds to a quasi-spherical surface surrounding the planet on which the mirror points must lie as the particle drifts in azimuth (i.e., local time).

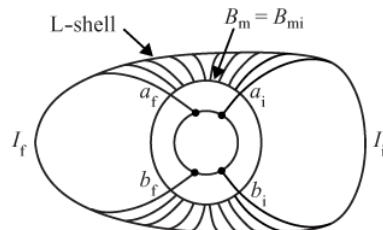


Fig. 3.26 As a particle drifts in azimuth in a planetary magnetic field it traces out a surface called an L-shell.

Gradient and curvature drifts cause a particle initially trapped on a field line (a_i, b_i) with integral invariant I_i to drift to a field line (a_f, b_f) that satisfies $I_f = I_i$. This field line can be identified by computing I for a series of field lines on the left side of the planet, starting at B_{mi} in the northern hemisphere and ending at B_{mi} in the southern hemisphere. This process is continued until the magnetic field line is found for which $I_f = I_i$. The locus of all field

lines around the planet with the same I for a given B_m generates a surface called an L-shell. As particles drift around the planet they stay on the same L-shell. The L -value for a given shell is determined by the equatorial radial distance measured in planetary radii of a field line with the same values of I and B_m in a dipole field that has the same magnetic moment as the planet. The L -value and mirror magnetic field B_m provide a very useful set of coordinates for ordering measurements of energetic charged particles trapped in planetary magnetic fields. For a further discussion, see McIlwain [1961].

3.6.2.3 The 3rd adiabatic invariant

The 3rd adiabatic invariant is associated with azimuthal drift motions and only exists for fields with a well-defined axial symmetry, such that the particles drift around in nearly closed orbits as illustrated in Figure 3.27. For such orbits, the adiabatic invariant is simply the magnetic flux linking the drift orbit. To prove this result, for simplicity we only consider orbits near the axis of symmetry. If the magnetic field changes with time, it produces an azimuthal non-conservative electric field given by Faraday's law

$$\int_C \mathbf{E} \cdot d\ell = - \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{A}, \quad (3.6.20)$$

where the integral is evaluated along the drift contour. Making use of the assumed axial symmetry and the weak radial dependence of B near the axis, the integrals in the above equation can be evaluated easily and show that

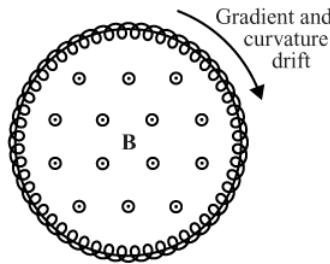


Fig. 3.27 In a slowly varying magnetic field, the magnetic flux through the drift orbit is constant.

$$2\pi R E = - \frac{dB}{dt} (\pi R^2) \quad (3.6.21)$$

or

$$E = - \frac{R}{2} \frac{dB}{dt}. \quad (3.6.22)$$

The azimuthal electric field produces a radial $\mathbf{E} \times \mathbf{B}$ drift, given by

$$v_E = \frac{E}{B} = - \frac{R}{2B} \frac{dB}{dt}. \quad (3.6.23)$$

Since $v_E = dR/dt$ it follows that

$$2\frac{dR}{R} = -\frac{dB}{B}. \quad (3.6.24)$$

The above equation can be integrated directly to show that the magnetic flux inside the drift orbit is constant, that is,

$$\Phi_B = \pi R^2 B = \text{constant}. \quad (3.6.25)$$

This is the 3rd adiabatic invariant. The 3rd adiabatic invariant shows that the guiding center of a particle drifting azimuthally in a slowly varying axially symmetric magnetic mirror field stays on the surface of a flux tube.

3.6.3 Violations of the adiabatic invariants

Exact conservation of the three adiabatic invariants implies that once a particle is trapped in an axially symmetric magnetic mirror field it should remain trapped forever, spiraling around the magnetic field, bouncing back and forth between the mirror points, and drifting azimuthally around the axis of symmetry. Because experiments with magnetic confinement systems often show significant deviations from this idealized description, it is important to consider how the adiabatic invariants can be violated.

Since the adiabatic invariants are all related to action integrals, it is easy to see that there are two ways in which the adiabatic invariants can be violated. First, the parameters of the system may vary so rapidly that the action integral is no longer an approximate constant of the motion. Second, the parameters may change in such a way that the system is no longer periodic, in which case the action integral no longer exists. Examples illustrating the various ways in which the three adiabatic invariants can be violated are discussed separately below.

3.6.3.1 Violations of the 1st invariant

The conditions for violation of the 1st adiabatic invariant are relatively easy to determine because this invariant is directly related to the WKB solution for the time varying harmonic oscillator. The condition for validity of the WKB solution is given by equation (3.6.9) with $\omega = \omega_c$ and is violated whenever either $\dot{\omega}_c$ or $\ddot{\omega}_c$ becomes sufficiently large, or whenever ω_c goes through zero. The variations can be either due to an explicit time dependence, or due to the motion of the guiding center through an inhomogeneous magnetic field.

One of the most important examples of an explicit time dependence arises from the presence of waves. Interactions with waves are particularly strong if the wave frequency, ω' , in the guiding center frame of reference is at the cyclotron frequency or a multiple of the cyclotron frequency

$$\omega' = \omega - k_{||}v_{||} = n\omega_c, \quad (3.6.26)$$

where n is a non-zero integer. This type of interaction is called a resonant interaction, and the above condition is called cyclotron resonance. The term $k_{\parallel}v_{\parallel}$, where \mathbf{k} is the wave vector, represents the Doppler shift. At cyclotron resonance, the wave field rotates in synchronism with the cyclotron motion of the particle. Even for very small wave intensities this type of interaction can cause large changes in the perpendicular kinetic energy, thereby violating the 1st adiabatic invariant.

Abrupt spatial changes in the magnetic field, such as at shock waves and other types of discontinuities, can also violate the 1st adiabatic invariant. These effects are particularly important for ions, since ions have much larger cyclotron radii than electrons and are more susceptible to spatial changes.

The 1st adiabatic invariant is also violated whenever the cyclotron frequency goes through zero. At zero frequency the motion is no longer periodic, so the action integral no longer exists. An important example where this type of violation occurs is near an x-type neutral line, where the magnetic field, and hence the cyclotron frequency, goes to zero ($B = 0$). It can be shown that the simplest magnetic field geometry near a neutral line consists of an x-shaped separatrix that separates regions of oppositely directed magnetic fields, such as shown in Figure 3.28. The neutral line in this diagram passes through the origin perpendicular to the plane of the paper. An x-type neutral line of this type is believed to form between the Earth's magnetic field and the solar wind [Dungey, 1961], which is a high speed plasma that flows radially outward from the sun [Hundhausen, 1972]. The geometry is easiest to picture when the magnetic field in the solar wind is directed southward, opposite to the Earth's magnetic field near the equator. Under these conditions the neutral line encircles the Earth. If a two-dimensional cut is made through the magnetosphere, such as in Figure 3.29, the neutral line appears as two x-type separatrices similar to Figure 3.28, one on the dayside (to the left in the diagram), and the other on the nightside (to the right) in a region called the magnetotail. Since the magnetic field reverses direction across the neutral line, particles moving in the vicinity of the neutral line follow S-shaped trajectories in the y,z plane as shown in Figure 3.30. For trajectories of this type, the 1st adiabatic invariant is obviously violated since the essential cyclotron motion no longer occurs. Because charges of opposite sign move in opposite directions, the S-type trajectories cause a current to flow along the neutral line. This current must be present to satisfy Ampère's law, $\nabla \times \mathbf{B} = \mu_0 \mathbf{J}$, in the region of oppositely directed field.

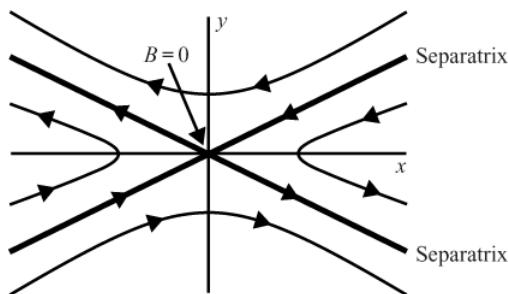


Fig. 3.28 The magnetic field geometry around an x-type neutral point.

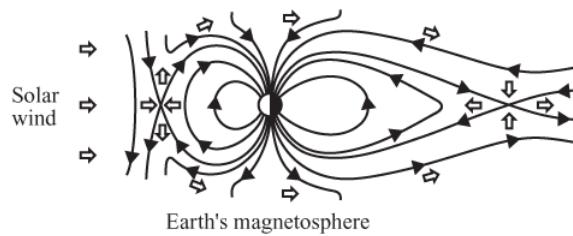


Fig. 3.29 The magnetic field configuration near the Earth when the solar wind magnetic field is directed southward, opposite to the Earth's magnetic field.

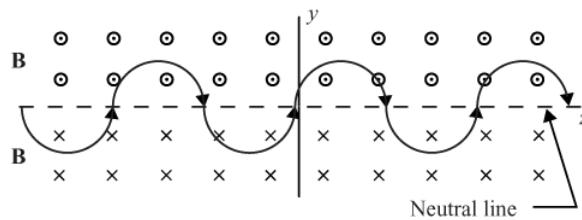


Fig. 3.30 Particles moving along the neutral line follow S-shaped trajectories, which violate the 1st adiabatic invariant.

A particularly interesting modification to the basic x-type magnetic field configuration arises if a static electric field is present parallel to the neutral line. As shown in Figure 3.31, this electric field causes an $\mathbf{E} \times \mathbf{B}$ drift, \mathbf{v}_E , toward the neutral line (downward for $y > 0$, and upward for $y < 0$). This $\mathbf{E} \times \mathbf{B}$ drift continuously transports particles into the region where the 1st adiabatic invariant is violated. The current produced by the S-type trajectories along the neutral line also results in a positive Joule heating, $\mathbf{E} \cdot \mathbf{J}$, indicating that electrical energy is being dissipated. Viewed in the x,y plane, the $\mathbf{E} \times \mathbf{B}$ drift transports plasma across the x-shaped separatrices, from regions A–A' to regions B–B', as shown in Figure 3.32. This type of plasma transport, across a boundary between regions with topologically different magnetic field configurations, is called magnetic merging or reconnection. The terminology comes from magnetospheric physics. On the dayside of the Earth, the solar wind carries magnetic field lines into the nose of the magnetosphere where they are “merged” with the Earth’s magnetic field lines, as indicated by the open arrows near the dayside x-point in Figure 3.29. These same magnetic field lines are carried over the polar caps by the motion of the solar wind and are then “reconnected” at the nightside x-point. For a further discussion of this field geometry and the associated plasma flow, see Kivelson and Russell [1995]. Neutral lines of this type are also believed to occur in the solar corona [Parker, 1963] and have been demonstrated to occur in laboratory plasma experiments [Priest and Forbes, 2000]. We will return to the very important topic of magnetic field line merging and reconnection in Chapter ??.

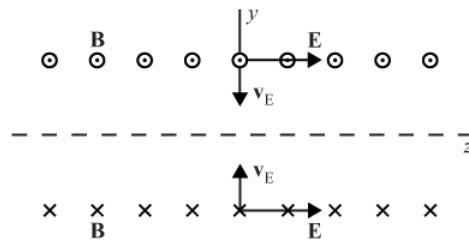


Fig. 3.31 If an electric field is imposed along the neutral line, the resulting $\mathbf{E} \times \mathbf{B}$ drift drives particles into the neutral line.

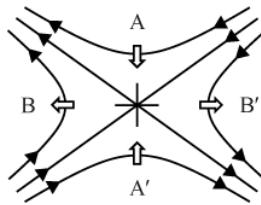


Fig. 3.32 The $\mathbf{E} \times \mathbf{B}$ drift velocities (large open arrows) near an x-type neutral line with a superimposed electric field directed upward, out of the paper.

3.6.3.2 Violations of the 2nd invariant

Electromagnetic perturbations that vary on time scales comparable to or less than the bounce time can violate the 2nd adiabatic invariant. Since the bounce time is usually much larger than the cyclotron period, the 2nd invariant is more easily violated than the 1st invariant. As with the 1st invariant, any type of resonance with the bounce motion provides a particularly effective mechanism for violating the 2nd invariant. A periodic interaction of this type is called a bounce resonance. Because the relevant motion occurs along the magnetic field line, the perturbing force must have a component parallel to the magnetic field. The main effect of a bounce resonance is to change the location of the mirror points, and therefore the total energy of the particle. Magnetohydrodynamic waves, discussed in Chapter 6, provide a common source of such perturbations.

If the magnetic field configuration changes to such an extent that the particle is no longer trapped, then the 2nd adiabatic invariant is violated because the corresponding action integral no longer exists. This can occur, for instance, when the mirror points are brought together by a compression parallel to the magnetic field.

Conservation of the 2nd adiabatic invariant

$$J = m \oint v_{\parallel} ds \quad (3.6.27)$$

shows that the parallel velocity, and therefore the total energy of the particle, must increase during such a compression. Since the magnetic mirror field strength required to reflect the particle increases in direct proportion to the total energy ($w = \mu B_m$), the particle energy eventually increases to the point that it can no longer remain trapped. The particle then es-

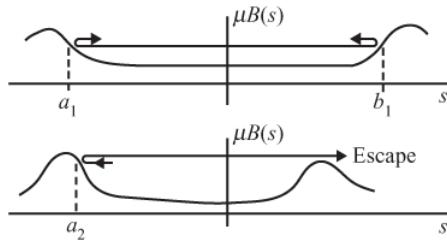


Fig. 3.33 Violation of the 2nd adiabatic invariant due to escape of the particle during a parallel compression.

capes from the system, violating the 2nd invariant. This type of process, illustrated in Figure 3.33, has been suggested by Fermi [1949] as a mechanism for accelerating cosmic rays to high energies. In Fermi's acceleration mechanism, a particle is reflected back and forth between regions of strong magnetic field embedded in two approaching plasma clouds. As the distance between the two approaching clouds decreases, the particle is accelerated by repeated reflection between the two clouds, conserving the 2nd adiabatic invariant. Eventually the particle energy increases to the point that the magnetic field in the clouds cannot reflect the particle, and the particle escapes. Fermi acceleration may also be effective in explaining the acceleration of particles by shock waves associated with solar flares and supernova explosions. As we will show in Chapter 7, a substantial jump in the magnetic field occurs at a shock wave. If a particle is trapped between the shock and a stationary reflection point ahead of the shock, substantial acceleration can occur before the particle escapes from the system.

3.6.3.3 Violations of the 3rd invariant

Since the azimuthal drift has the largest period of the three types of periodic motions that occur in a magnetic mirror, the 3rd invariant is the easiest to violate. Any change in the system that occurs on a time scale comparable to or shorter than the azimuthal drift period violates the 3rd invariant. In planetary magnetospheres, changes in electric and magnetic fields due to variations in the solar wind pressure can easily cause violations of the 3rd invariant. If the 3rd adiabatic invariant is violated, particles can move from one L-shell to another as they drift azimuthally around the planet. If the perturbing forces are random, the resulting radial motion can be described as a random walk, and is often called radial diffusion. Radial diffusion is believed to be the primary mechanism by which particles are transported radially in planetary radiation belts. If a particle diffuses into a dipole magnetic field from a large distance while conserving the 1st ($\mu = w_\perp/B$) and 2nd ($J = m \int v_\parallel ds$) adiabatic invariants, considerable energization can occur. Since the magnetic field strength of a dipole field varies as $1/R^3$, conservation of the first adiabatic invariant implies that the perpendicular energy should vary as $1/R^3$. This process is illustrated in Figure 3.34. As a particle moves from L_1 to L_2 , the perpendicular kinetic energy increases by a factor of $(L_2/L_1)^3$. Since the length of the magnetic field decreases substantially in going from L_1 to L_2 , conservation of the 2nd invariant shows that the parallel kinetic energy also increases

by a substantial factor. The change in the parallel energy is somewhat difficult to analyze since the mirror points change as the particle moves inward.

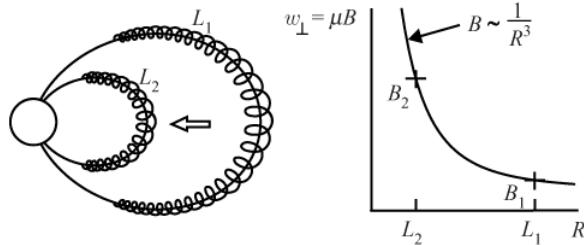


Fig. 3.34 Violations of the 3rd adiabatic invariant can cause particles to drift azimuthally onto different L-shells (L_1 to L_2). If the 1st adiabatic invariant is conserved, the increase in the magnetic field strength (B_1 to B_2) causes a considerable increase in the perpendicular energy, w_{\perp} .

3.7 The Hamiltonian method

Up to this point our treatment of charged particle motions in inhomogeneous and time varying magnetic fields has centered on certain adiabatic invariants that are valid in the limit of small spatial gradients and slow temporal variations. As is well known from classical mechanics, if a dynamical system has some geometric symmetry, then there is a corresponding quantity that is an exact constant of the motion. The analysis of such systems is best approached by using Hamilton's equations of motion [see Goldstein, 1959]. Although the Hamiltonian approach relies on the presence of symmetry, it does not assume adiabaticity, and thus complements the approach developed in Section 3.6.

3.7.1 Hamilton's equations

In Hamiltonian mechanics, the state of a system is specified by a set of generalized coordinates q_i and their conjugate momenta p_i . The Hamiltonian is defined by

$$H = \sum_i \dot{q}_i p_i - L, \quad (3.7.1)$$

where L is the Lagrangian. For a charged particle moving in an electromagnetic field the Lagrangian is given by

$$L = \frac{1}{2}mv^2 + q\mathbf{A} \cdot \mathbf{v} - q\Phi, \quad (3.7.2)$$

where \mathbf{A} is a vector potential defined by $\mathbf{B} = \nabla \times \mathbf{A}$, and Φ is a scalar potential defined by

$\mathbf{E} = -\nabla\Phi - \partial\mathbf{A}/\partial t$. The conjugate momenta p_i are obtained from the Lagrangian by the relation

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \quad (3.7.3)$$

Once the Hamiltonian for the system is known as a function of the generalized coordinates and momenta, the motion of the system can be determined from Hamilton's equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}, \quad \text{and} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \quad (3.7.4)$$

For a system with f degrees of freedom ($i = 1, \dots, f$), Hamilton's equations constitute a system of $2f$ first-order differential equations that can be solved for the $2f$ variables q_i and p_i .

To use Hamilton's equations it is first necessary to determine the vector and scalar potentials from the known electric and magnetic fields. This step involves a certain arbitrariness (known as gauge freedom), since $\nabla \cdot \mathbf{A}$ is not specified. Usually the choice of \mathbf{A} is made in such a way as to minimize the number of generalized coordinates by taking advantage of various symmetries. Once \mathbf{A} and Φ are determined, the conjugate momenta are computed using equation (3.7.3). The equations for the conjugate momenta are then used to eliminate the \dot{q}_i from equation (3.7.1), thus determining the Hamiltonian as a function of the coordinates, conjugate momenta and time, $H(q_i, p_i, t)$. Hamilton's canonical equations (3.7.4) then yield a system of non-linear differential equations that describe the motion of the system.

Although Hamilton's equations can in principle be solved to give the motion of the system, in most cases it is difficult, if not impossible, to obtain closed-form analytic solutions. Hamilton's equations are often non-integrable and sometimes exhibit "chaotic" solutions, that is, solutions that are so sensitive to initial conditions that extremely small differences in initial conditions produce exponentially large differences in the final solutions. However, if one or more constants of the motion can be identified, it is sometimes possible to obtain a qualitative understanding of the motions involved. If the Hamiltonian does not depend explicitly on a particular coordinate q_j , then from equation (3.7.4) one obtains

$$\dot{p}_j = -\frac{\partial H}{\partial q_j} = 0. \quad (3.7.5)$$

The corresponding conjugate momentum p_j is then a constant of the motion. Such a coordinate is called a cyclic coordinate. If, in addition, the Lagrangian is independent of time (i.e., static electric and magnetic fields) then

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} = 0, \quad (3.7.6)$$

which implies that the Hamiltonian itself is a constant. The Hamiltonian then corresponds to the total energy, W , of the system.

3.7.2 Hamiltonian for an axially symmetric magnetic mirror field

To illustrate the techniques involved in the Hamiltonian method we next analyze an axially symmetric magnetic mirror field. The results obtained can then be compared and contrasted with the results from the adiabatic invariant approach.

For an axially symmetric system it is natural to use cylindrical coordinates (ρ, ϕ, z) . The Lagrangian in cylindrical coordinates is

$$L = \frac{1}{2}m(\dot{\rho}^2 + \rho^2\dot{\phi}^2 + \dot{z}^2) + q(A_\rho\dot{\rho} + A_\phi\rho\dot{\phi} + A_z\dot{z}), \quad (3.7.7)$$

and the conjugate momenta are

$$p_\rho = m\dot{\rho} + qA_\rho, \quad (3.7.8)$$

$$p_\phi = m\rho^2\dot{\phi} + q\rho A_\phi, \quad \text{and} \quad (3.7.9)$$

$$p_z = m\dot{z} + qA_z. \quad (3.7.10)$$

After eliminating $\dot{\rho}$, $\dot{\phi}$, and \dot{z} from equation (3.7.7), the Hamiltonian can be written

$$H = \frac{1}{2m}(p_\rho - qA_\rho)^2 + \frac{(p_\phi - q\rho A_\phi)^2}{2m\rho^2} + \frac{1}{2m}(p_z - qA_z)^2. \quad (3.7.11)$$

If the magnetic field has axial symmetry with no B_ϕ component, the field can be derived from a vector potential of the form $\mathbf{A} = [0, A_\phi(\rho, z), 0]$ with

$$B_\rho = -\frac{\partial A_\phi}{\partial z} \quad \text{and} \quad B_z = \frac{1}{\rho}\frac{\partial}{\partial \rho}(qA_\phi). \quad (3.7.12)$$

The conjugate momenta are then

$$p_\rho = m\dot{\rho}, \quad p_\phi = m\rho^2\dot{\phi} + q\rho A_\phi, \quad \text{and} \quad p_z = m\dot{z}. \quad (3.7.13)$$

Since ϕ is a cyclic coordinate ($\partial H/\partial\phi = 0$), from equation (3.7.5) one can see that

$$p_\phi = \text{constant}. \quad (3.7.14)$$

If the magnetic field is time independent, the Hamiltonian is also a constant of the motion, so that

$$H = \frac{p_\rho^2}{2m} + \frac{p_z^2}{2m} + \psi(\rho, z) = W, \quad (3.7.15)$$

where $\psi(\rho, z)$ is an effective potential defined by

$$\psi(\rho, z) = \frac{1}{2m\rho^2} (p_\phi - q\rho A_\phi)^2, \quad (3.7.16)$$

and W is a constant.

When the Hamiltonian is written in the form (3.7.15) it is evident that the motion is equivalent to the two-dimensional motion of a particle of mass m and total energy W in a potential $\psi(\rho, z)$. The motion is then bounded by the surface

$$W = \psi(\rho, z). \quad (3.7.17)$$

All that remains is to determine the general form of $\psi(\rho, z)$. Since p_ϕ is determined by the initial conditions, it is obvious that some value for p_ϕ exists such that $p_\phi = q\rho A_\phi$. The effective potential (at a fixed value of z) then has a minimum at a ρ -value determined by the condition $p_\phi = q\rho A_\phi$. A representative plot of $\psi(\rho, z)$ for a constant value of z is shown in Figure 3.35. If the total energy W is sufficiently small, one can see that the motion along the ρ direction is always bounded, with one bound inside and the other bound outside the point where $p_\phi = q\rho A_\phi$.

The location of the minimum at $p_\phi = q\rho A_\phi$, has an interesting interpretation. Let us compute the magnetic flux Φ_B through a surface of radius ρ perpendicular to the z axis

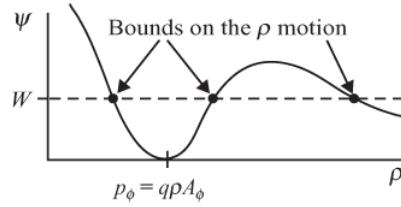


Fig. 3.35 The effective potential, ψ , and the bounds on the ρ motion.

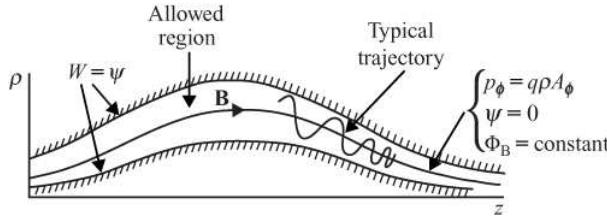


Fig. 3.36 The bounds on the motion determined by the condition $W = \psi(\rho, z)$.

$$\Phi_B = \int_0^\rho B_z 2\pi\rho d\rho = \int_0^\rho \left[\frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho A_\phi) \right] 2\pi\rho d\rho. \quad (3.7.18)$$

The integral can be evaluated directly to give

$$\Phi_B = 2\pi\rho A_\phi. \quad (3.7.19)$$

Thus, the condition

$$P_\phi = q\rho A_\phi(\rho, z) = \text{constant} \quad (3.7.20)$$

corresponds to a surface of a flux tube (i.e., a magnetic field line). Therefore, in the ρ, z plane, the effective potential has a minimum that follows a magnetic field line. As shown in Figure 3.36, the bounds on the motion prescribed by equation (3.7.17) represent two surfaces on either side of this magnetic field line. A particle moving in this two-dimensional potential well executes an oscillatory motion in the ρ direction while moving in the z direction, always remaining within the bounds given by $W = \psi(\rho, z)$. A turning point in the z motion occurs as the potential well narrows in the region of strong magnetic field. This turning point corresponds to the mirror point described in Section 3.6 using adiabatic invariants.

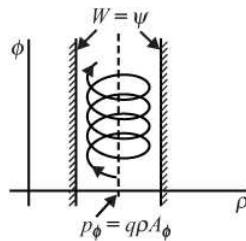


Fig. 3.37 The azimuthal motion.

The azimuthal motion can be determined from the equation for p_ϕ , see (3.7.13), which can be solved for $\dot{\phi}$ to give

$$\dot{\phi} = \frac{p_\phi - q\rho A_\phi}{mp^2}. \quad (3.7.21)$$

It is obvious that $\dot{\phi}$ changes sign every time the particle crosses the magnetic field line (where $p_\phi = q\rho A_\phi$). Therefore, the azimuthal motion consists of loops in the ρ, ϕ plane, as shown in Figure 3.37. The looping motion corresponds to the usual cyclotron motion around the magnetic field line. The asymmetry of the loops causes a secular motion in the ϕ direction that corresponds to the gradient drift discussed in Section 3.3.1. If the particle has a total energy sufficiently small to be trapped near the bottom of the potential well (see Figure 3.35), where the potential is parabolic, then the ρ and ϕ motions are harmonic, corresponding to the usual nearly circular cyclotron motion.

If the magnetic field varies with time, then the total energy W is no longer constant and the bounds on the motion given by $W = \psi$ change with time. The variations could cause the particle to escape from the potential well. However, if the particle does not gain enough energy to escape from the potential well, it will continue to oscillate around a field line with the same constant magnetic flux Φ_B . This result is essentially equivalent to the conservation of the 3rd adiabatic invariant, but generalized to allow an arbitrary radial dependence for the magnetic field.

It is useful to compare the results of the Hamiltonian and adiabatic invariant methods

of analysis. Both methods show that in an axially symmetric static magnetic mirror field the particle performs a cyclotron motion around a magnetic field line while simultaneously bouncing back and forth between regions of strong magnetic field and drifting azimuthally around the axis of symmetry. However, the Hamiltonian method is more robust. The Hamiltonian method requires no assumption about the amplitude of the oscillations or the magnitude of the spatial gradients in the magnetic field. Furthermore, if the magnetic field is time independent, the Hamiltonian method in some cases can place a rigorous bound on the particle motions.

3.7.3 Motion in a dipole magnetic field

The motion of a charged particle in a dipole magnetic field is a basic problem that has applications to the study of both cosmic rays and planetary radiation belts. The general nature of the motion was first solved in a classic paper by Störmer [1907]. Because a dipole field has axial symmetry, the Hamiltonian method is ideally suited to this problem. However, instead of the cylindrical coordinates used in Section 3.7.2, it is more convenient to use spherical coordinates r, θ, ϕ , where r is the radius, θ is the polar angle, and ϕ the azimuthal angle.

An appropriate vector potential for a dipole magnetic field in spherical coordinates is

$$A_\phi = \frac{\mu_0 M}{4\pi} \frac{\sin \theta}{r^2}, \quad (3.7.22)$$

where M is the magnetic moment. The effective potential is easily shown to be

$$\psi = \frac{1}{2mr^2 \sin^2 \theta} \left(p_\phi - \frac{q\mu_0 M \sin^2 \theta}{4\pi r} \right)^2. \quad (3.7.23)$$

The minimum in the effective potential, corresponding to $\psi = 0$, is located at

$$r = \frac{q\mu_0 M}{4\pi p_\phi} \sin^2 \theta, \quad (3.7.24)$$

which is easily verified to be the equation for a dipole magnetic field line in spherical coordinates. The bounds on the motion are obtained by setting $\psi = W$, which gives

$$W = \frac{1}{2mr^2 \sin^2 \theta} \left(p_\phi - \frac{q\mu_0 M \sin^2 \theta}{4\pi r} \right)^2. \quad (3.7.25)$$

The above equation has solutions

$$\frac{1}{r} = \frac{2\pi p_\phi}{q\mu_0 M \sin^2 \theta} \left[1 \pm \sqrt{1 \pm \frac{q\mu_0 M}{\pi p_\phi^2} \sqrt{2mW} \sin^3 \theta} \right]. \quad (3.7.26)$$

If p_ϕ is negative only one root has a real positive value, whereas if p_ϕ is positive either one or three roots have real positive values. The number of roots when p_ϕ is positive is controlled by the quantity under the square root, which must be positive to give real values

for r . The bounds on the motion can be summarized by the three diagrams in Figure 3.38. Note that at an infinite distance the quantity p_ϕ is just the angular momentum, which can be written

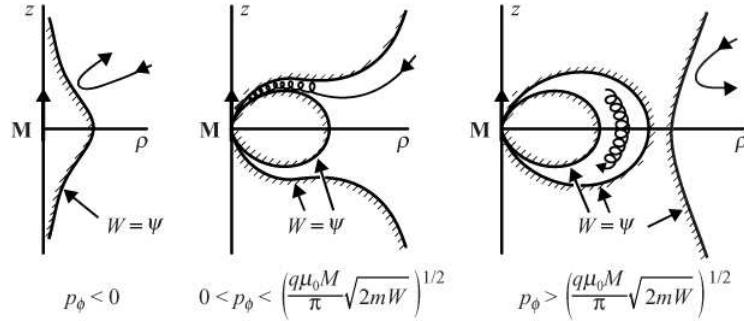


Fig. 3.38 The bounds on the motion for a charged particle moving in a dipole magnetic field.

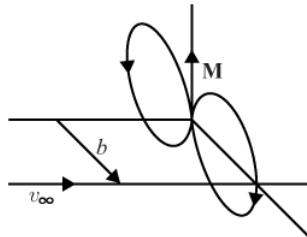


Fig. 3.39 The geometry defining the impact parameter, b .

$$p_\phi(\rho = \infty) = mbv_\infty, \quad (3.7.27)$$

where b is the impact parameter relative to the z axis and v_∞ is the speed at infinity, as shown in Figure 3.39. For negative p_ϕ , all particles are reflected away from the dipole, irrespective of their total energy, as illustrated in the left diagram of Figure 3.38. This occurs because the $\mathbf{v} \times \mathbf{B}$ force has a component directed outward, away from the dipole. For positive p_ϕ , the $\mathbf{v} \times \mathbf{B}$ force has a component of force directed inward, toward the dipole, so the trajectories curve toward the dipole. For certain initial conditions the particles can approach the dipole very closely by going through the “throat” region near the equator and then spiraling down a magnetic field line, as illustrated in the middle diagram of Figure 3.38.

For p_ϕ greater than a critical value, given by

$$p_\phi > \left(\frac{q\mu_0 M}{\pi} \sqrt{2mW}\right)^{1/2}, \quad (3.7.28)$$

the allowed region breaks into two distinct regions, as illustrated in the right diagram of Figure 3.38. One of these regions corresponds to internally trapped particles, and accounts

for the radiation belts of the Earth and other planets [Van Allen, 1996]. It is interesting to note that the Hamiltonian method shows that in the absence of time variations this region is rigorously inaccessible to particles arriving from the outside, a result that cannot be claimed from the adiabatic invariant approach. However, even with the Hamiltonian method, the motion of a charged particle in a dipole magnetic field falls into a class of problems that are “insoluble” in the formal sense [Dragt and Finn, 1976], and can lead to chaotic orbits, which is the topic of the next section.

3.8 Chaotic orbits

In the previous discussion we have shown several examples of particle motion where, either for reasons of symmetry or because the particle Hamiltonian depends on parameters that vary slowly with time (or space), the motion of the particle is constrained either by exactly conserved quantities or by adiabatic invariants. For example, in Section 3.7.2, we considered the motion of a charged particle in a time independent and axially symmetric magnetic field. Time independence implies time translation symmetry ($\partial/\partial t = 0$), and the corresponding conserved quantity is the Hamiltonian given by equation (3.7.15) which, in this case, is the total energy, W . Similarly, axial symmetry ($\partial/\partial\phi = 0$) implies that the momentum p_ϕ , given by equation (3.7.13), is a constant. The existence of these two conserved quantities, W and p_ϕ , drastically constrains the motion of the particle, as indicated in Figures 3.36 and 3.37. Even in the absence of rigorous symmetries, if the Hamiltonian depends on parameters that vary slowly, adiabatic invariants play the role of constants of the motion.

We have discussed in Section 3.6.2 how the three adiabatic invariants of particle motion in a magnetic mirror geometry may be broken if certain parameters vary rapidly. Constants of the motion, such as W and p_ϕ , can also be destroyed by perturbations that break the symmetries in t and ϕ . When such violations or breakdowns occur, either because existing symmetries are broken or because certain slowly varying parameters vary too rapidly, the particle motion is not constrained by either exact or adiabatic invariants and can become chaotic. Chaotic motions are characterized by an extreme sensitivity to the initial conditions. Typically, the space made up of the initial conditions has regions that lead to bounded integrable orbits, and other regions where the orbits are chaotic. It is now known that chaotic motions are a generic feature of Hamiltonian systems, and their occurrence is an important feature that must be considered when analyzing charged particle motions in plasmas.

Sometimes even very simple magnetic field geometries lead to chaotic motions. One of the simplest is a two-dimensional magnetic field containing an x-type neutral point, such as shown in Figure 3.28. Since the 1st adiabatic invariant is always violated in the vicinity of the neutral point, the trajectories for such geometries must be determined by numerical computations. An example of such a trajectory, from Chen and Palmaresso [1986], is shown in Figure 3.40. The S-shaped trajectory, similar to that expected near an x-type neutral line (see Figure 3.30), is clearly evident near $x = 0$. In their studies of trajectories

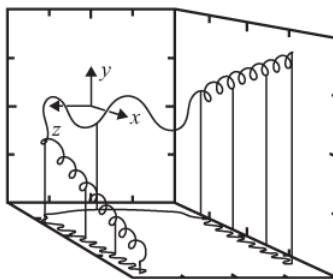


Fig. 3.40 A representative particle trajectory near an x-type neutral line using the same coordinate system as in Figure 3.28 [adapted from Chen and Palmadesso, 1986].

of this type, Chen and Palmadesso [1986] and also Büchner and Zelenyi [1989] showed that, depending on the initial conditions, certain particles displayed chaotic behavior, i.e., trajectories that diverged exponentially from trajectories with only slightly different initial conditions. The distinction between integrable and chaotic orbits is particularly clear for a parabolic magnetic field configuration of the form $\mathbf{B} = B_0[\hat{\mathbf{y}} + \hat{\mathbf{x}}(y/L)]$, where B_0 and L are constants. Such a field configuration, which is shown in Figure 3.41, is qualitatively similar to the field that exists along the current sheet near an x-type neutral point. It can be shown that the length scale, L , corresponds to the radius of curvature of the magnetic field line at $y = 0$. Since the magnetic field strength increases with distance both above and below the symmetry plane, $y = 0$, all of the particles are magnetically trapped, irrespective of their pitch angle at $y = 0$. At first glance it appears that both the x and z coordinates are cyclic, since the magnetic field is independent of these two coordinates. If this were the case, there would be three constants of the motion, the total energy, W , the x component of the conjugate momentum, p_x , and the z component of the conjugate momentum, p_z . The trajectories would then be fully integrable, since there are as many constants of the motion as there are generalized coordinates. However, this is not the case, since what counts in the Hamiltonian is the vector potential, and the vector potential for this field unavoidably has either an x or z dependence. For example, if the field is represented by the vector potential $\mathbf{A} = \hat{\mathbf{z}}B_0(-x + y^2/2L)$, the z coordinate is cyclic, but the x coordinate is not. For purposes of plotting the trajectory, it is useful to use the normalized variable $X = (x - p_z/qB_0)/L$, where $p_z = \text{constant}$. As the particle moves in the magnetic field, it repeatedly crosses the symmetry plane ($y = 0$), while conserving the total energy, W . To reveal possible chaotic orbits, a phase-space plot is constructed using the X and \dot{X} coordinates each time a particle crosses the $y = 0$ plane. Such a plot is shown in Figure 3.42. This type of plot is called a Poincaré surface of section, after the French mathematician Poincaré who first used this type of plot to study planetary orbits [Gleick, 1987]. In this plot, integrable orbits appear as the concentric closed “curves” in region A, each of which is made up of successive $y = 0$ crossing points. In contrast, a chaotic orbit leaves seemingly “random” crossing points, as in region B. No orbit in region A crosses into region B and vice versa. The outermost solid circle corresponds to an orbit with zero parallel velocity in the $y = 0$ plane, i.e., a trajectory consisting of a simple cyclotron orbit in the $y = 0$ plane. For the example shown, the ratio of the cyclotron radius, ρ_c , of this orbit to the radius of curvature, L , is $\rho_c/L = 31.6$. The

transition between the integrable region and the chaotic region is not sharp, and has fractal-like structures extending into the chaotic region, as often occurs in such plots. Although the trajectories in region B appear random they are not, since each point is determined by the initial conditions. However, even the smallest random fluctuation in the initial conditions, or the magnetic field (i.e., noise), leads to a random distribution. Such chaotic orbits are now known to be an important feature of particle motions near x-type neutral lines in planetary magnetospheres, and in certain laboratory magnetic confinement machines, such as the reverse field pinch (see Chapter ??).

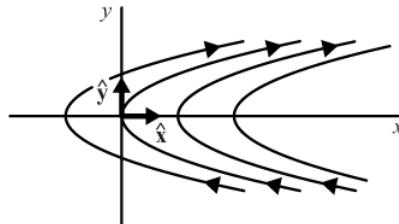


Fig. 3.41 A parabolic magnetic field configuration, such as exists along the current sheet near an x-type neutral point (compare with Figure 3.28).

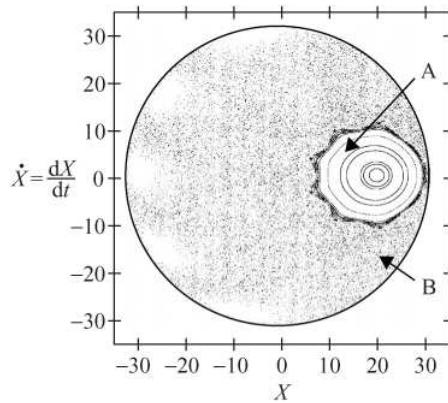


Fig. 3.42 The normalized position and velocity coordinates (X, \dot{X}) at which particles cross the $y = 0$ plane for a parabolic magnetic field geometry, such as occurs near an x-type neutral line [adapted from Chen and Palmaresso, 1986]. The particles in region A are in bounded integrable orbits, and the particles in region B have chaotic orbits.

Chaotic orbits can also arise from the structure of the magnetic field. If the symmetry of the system is suitably broken [for examples, see Lau and Finn, 1991; Priest *et al.*, 1996], the spatial configuration of the magnetic field lines may then become chaotic, i.e., adjacent magnetic field lines diverge exponentially from each other, no matter how small the initial separation. For such chaotic magnetic field configurations, the motion of the particle is chaotic, even if the guiding center of a particle follows a magnetic field line exactly. For

a further discussion of such chaotic magnetic field configurations see Priest and Forbes [2000].

Problems

- 3.1. A plasma has a constant uniform magnetic field aligned along the z axis, $\mathbf{B} = B_0 \hat{\mathbf{z}}$, and has a constant uniform drift along the x axis, $\mathbf{U} = v_E \hat{\mathbf{x}}$. A positively charged particle of charge $+q$ and mass m is released from rest at the origin.
- (a) Compute the electric field in the plasma.
 - (b) Find the equation of motion of the particle (i.e., $x(t)$, $y(t)$, and $z(t)$).
 - (c) What is the kinetic energy of the particle in the plasma rest frame?
- 3.2. A cylindrical column of plasma rotates around its central axis (as though it were a rigid solid) at an angular velocity ω_0 . A constant uniform magnetic field \mathbf{B}_0 is present parallel to the axis of rotation.
- (a) Assuming that the rigid rotation can be described by $\mathbf{v}_E = \omega_0 \times \mathbf{r}$, where $\mathbf{v}_E = \mathbf{E} \times \mathbf{B}/B^2$, compute the electric field \mathbf{E} in the plasma column. Use cylindrical (ρ, ϕ, z) coordinates.
 - (b) Is there a polarization charge $\rho_q = \epsilon_0 \nabla \cdot \mathbf{E}$ associated with this electric field? If so, how does ρ_q depend on the distance from the central axis?
 - (c) Find the electrostatic potential ϕ such that $\mathbf{E} = -\nabla\phi$.
 - (d) How would you induce a motion of this type in a magnetized plasma column?
- 3.3. A planetary magnetosphere has a dipole magnetic field with a dipole moment \mathbf{M} . The dipole axis is aligned along the rotational axis of the planet. All of the plasma in the magnetosphere rotates with the planet at an angular velocity ω . Using spherical (r, θ, ϕ) coordinates derive equations for the following.
- (a) The magnetic field, \mathbf{B} .
 - (b) The $\mathbf{E} \times \mathbf{B}$ drift velocity, $\mathbf{v}_E = \omega \times \mathbf{r}$.
 - (c) The electric field \mathbf{E} .
 - (d) The electrostatic potential, ϕ , such that $\mathbf{E} = -\nabla\phi$.
- 3.4. The equation for a dipole magnetic field in spherical coordinates is given by
- $$\mathbf{B} = \frac{\mu_0 M}{4\pi r^3} [\hat{\mathbf{r}} 2 \cos \theta + \hat{\theta} \sin \theta],$$
- where M is the magnetic moment.
- (a) Show that the equation for a magnetic field line is $r = R \sin^2 \theta$, where R is the radius of the magnetic field line at the equator ($\theta = \pi/2$).
 - (b) Show that the curvature of the magnetic field line at the equator ($\theta = \pi/2$) is $R_C = R/3$.
 - (c) Compute the curvature drift of a particle with charge q and parallel energy $w_{||}$ at a radial distance R at the equator.
 - (d) Compute the gradient drift of a particle with charge q and perpendicular energy w_{\perp} at a radial distance R at the equator.
 - (e) Compare the equations for the curvature drift and gradient drift at the equator.
- 3.5. A charged particle of mass m and charge q has its guiding center at a radius ρ from

the center of a long straight wire carrying a current I_0 . A uniform electric field E_0 exists parallel to the axis of the wire.

(a) Using cylindrical coordinates (ρ, ϕ, z) , derive equations for the $\mathbf{E} \times \mathbf{B}$ drift velocity, the gradient drift velocity, and the curvature drift velocity in terms of the quantities I_0 and E_0 .

(b) Solve for the ρ component of the guiding center motion as a function of time.

3.6. The ratio of the maximum to minimum magnetic field in a magnetic mirror machine is called the mirror ratio, $r = B_{\max}/B_0$. Assume that the particles initially have an isotropic velocity distribution at B_0 . What fraction of the particles at B_0 is trapped in the machine?

3.7. A particle is trapped in a magnetic mirror field given by

$$B_z = B_0 \left[1 + \left(\frac{z}{L} \right)^2 \right],$$

and has a total kinetic energy w and pitch angle α at $z = 0$. Find the oscillation frequency in terms of L , w , and α .

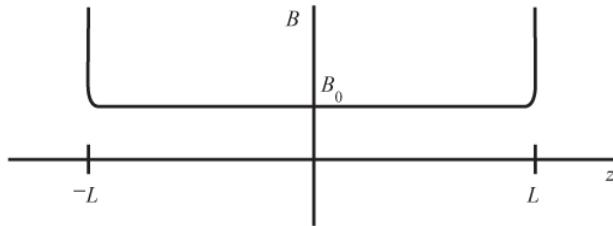
3.8. Demonstrate by direct calculation that the WKB solution

$$X_{\text{WKB}} = \frac{1}{\sqrt{\omega(t)}} \exp \left[\pm i \int^t \omega(t') dt' \right]$$

satisfies the differential equation

$$\frac{d^2X}{dt^2} + \left[\omega^2 + \frac{1}{2} \left(\frac{\ddot{\omega}}{\omega} \right) - \frac{3}{4} \left(\frac{\dot{\omega}}{\omega} \right)^2 \right] X = 0.$$

3.9. A particle is trapped in an axially symmetric magnetic mirror field. The field has a uniform value B_0 in the region $-L < z < L$, and increases to a very large value at $z = \pm L$ as shown below.



The initial pitch angle of a particle at $z = 0$ is $\alpha_0 = 45^\circ$. The ends of the mirror are now slowly moved into $\pm L/2$ while holding B_0 constant. Assume that the 1st and 2nd adiabatic invariants are constant. By what factor (w_f/w_i) does the total energy of the particle increase during this process?

3.10. A particle is trapped in a magnetic mirror field given by

$$B_z = B_0 \left[1 + \left(\frac{z}{L} \right)^2 \right].$$

Initially, the mirror points of the particle are located at $z = \pm L$.

- (a) B_0 is now slowly increased to $2B_0$. Using the 2nd adiabatic invariant, find the new mirror point locations and the new mirror field B_m .
- (b) L is then slowly decreased to $L/2$, while holding $2B_0$ constant. Using the 2nd adiabatic invariant, find the new mirror point locations and the new mirror field B_m .

References

- Brillouin, L. 1926. La mécanique ondulatoire de Schrödinger, une méthode générale de résolution par approximations successives. *C. R. Acad. Sci.* **183**, 24–26.
- Büchner, J., and Zelenyi, L. M. 1989. Regular and chaotic charged particle motion in magnetotail like field reversals. 1. Basic theory of trapped motion. *J. Geophys. Res.* **94**, 11821–11842.
- Chen, J., and Palmadesso, P. J. 1986. Chaos and nonlinear dynamics of single-particle orbits in a magnetotail-like magnetic field. *J. Geophys. Res.* **91**, 1499–1508.
- Dragt, A. J., and Finn, J. M. 1976. Insolubility of trapped particle motion in a magnetic dipole field. *J. Geophys. Res.* **81**, 2327–2340.
- Dungey, J. W. 1961. Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett.* **6**, 47–48.
- Fermi, E. 1949. On the origin of the cosmic radiation. *Phys. Rev.* **75**, 1169–1174.
- Gleick, J. 1987. *Chaos*. Penguin: New York, pp. 142–144.
- Goldstein, H. 1959. *Classical Mechanics*. Addison-Wesley: Reading, MA, p. 217.
- Hundhausen, A. J. 1972. *Coronal Expansion and the Solar Wind*. Springer-Verlag: New York, p. 17.
- Kivelson, M. G., and Russell, C. T. 1995. *Introduction to Space Physics*. Cambridge University Press: Cambridge, p. 243.
- Kramers, H. A. 1926. Wellenmechanik und halbzahlige Quantisierung. *Z. Phys.* **39**, 828–840.
- Lau, Y.-T., and Finn, J. M. 1991. Three-dimensional kinematic reconnection of plasmoids. *Astrophys. J.* **366**, 577–599.
- McIlwain, C. E. 1961. Coordinates for mapping the distribution of magnetically trapped particles. *J. Geophys. Res.* **66**, 3681–3691.
- Parker, E. N. 1963. The solar-flare phenomenon and the theory of reconnection and annihilation of magnetic fields. *Astrophys. J. Suppl.* **8**, 177–211.
- Priest, E. R., and Forbes, T. G. 2000. *Magnetic Reconnection: MHD Theory and Applications*. Cambridge University Press: Cambridge, p. 304.
- Priest, E. R., Lonie, D. P., and Titov, V. S. 1996. Bifurcations of magnetic topology by the creation or annihilation of null points. *J. Plasma Phys.* **56**, 507–530.
- Störmer, C. 1907. Sur des trajectoires des corpuscules électrisés dans l'espace sous l'action du magnétisme terrestre, Chapitre 4. *Arch. Sci. Phys. Nat.* **24**, 317–264.
- Van Allen, J. A. 1996. Kuiper prize lecture: Electrons, protons, and planets. *Icarus* **122**, 209–232.
- Wentzel, G. 1926. Eine Verallgemeinerung der Quantenbedingung für die Zwecke der Wellenmechanik. *Z. Phys.* **38**, 518–529.

Further reading

Baumjohann, W., and Treumann, R. A. *Basic Space Plasma Physics*. Imperial College Press: London, 1997, Chapter 2.

Chen, F. F. *Introduction to Plasma Physics and Controlled Fusion*, Vol. 1: *Plasma Physics*. Plenum Press: New York, 1983 (reprinted in 1990), Chapter 2. Northrop, T. G. *The Adiabatic Motion of Charged Particles*. Interscience: New York, 1963, Chapters 2 and 3.

Parks, G. K. *Physics of Space Plasmas: An Introduction*. Addison-Wesley: Redwood City, CA, 1991 (reprinted in 2000), Chapter 4.

Schmidt, G. *Physics of High Temperature Plasmas: An Introduction*. Academic Press: New York, 1966 (reprinted in 1979), Chapter 2.

Next we discuss the dispersion relation for a Maxwellian distribution function, which in normalized one-dimensional form is given by

Due to the presence of long-range forces, various types of waves can exist in a plasma that have no counterpart in ordinary gases or dielectric media. In this chapter we consider the special case of small amplitude waves propagating in a cold plasma. The assumption that the amplitude is small greatly simplifies the analysis by allowing the equations of motion to be linearized. The assumption that the plasma is cold means that the particles are initially at rest, with no thermal motions, thereby allowing us to ignore Doppler shifts, which provides a considerable simplification.

Since a cold plasma has no pressure, waves that depend on pressure, such as sound waves, cannot occur in a cold plasma. Also, since there is no source of free energy, instabilities are absent. Despite these limitations, the cold plasma model gives a remarkably accurate description of many types of waves that occur in real plasmas.

4.1 Fourier representation of waves

For reasons that will be apparent shortly, it is convenient to use Fourier transforms for solving linearized wave problems. Under very general conditions it can be shown that a function $f(x)$ can be represented by the integral

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} dk, \quad (4.1.1)$$

where $\tilde{f}(k)$ is called the Fourier transform of $f(x)$. The Fourier transform can be determined by inverting the above integral, and is given by

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx. \quad (4.1.2)$$

The functions $f(x)$ and $\tilde{f}(k)$ are often referred to as a Fourier transform pair, and are tabulated in mathematics handbooks for most commonly used functions. If x represents a spatial coordinate then $k = 2\pi/\lambda$ is called the wave number, where k is the wavelength.

A specific example that illustrates many of the general properties of Fourier transforms is the Gaussian harmonic packet,

$$f(x) = e^{-a_0^2 x^2/2} e^{ik_0 x}. \quad (4.1.3)$$

The parameter a_0 controls the width of the Gaussian envelope and $k_0 = 2\pi/\lambda_0$ is the wave

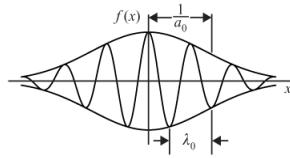


Fig. 4.1 The Gaussian harmonic wave packet.

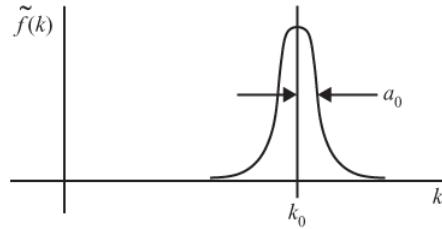


Fig. 4.2 The Fourier transform of the Gaussian harmonic wave packet.

number of the harmonic term. The real part of $f(x)$ is shown in Figure 4.1. It is straightforward to show that the Fourier transform of the Gaussian harmonic packet is given by

$$\tilde{f}(k) = \frac{1}{a_0} e^{-(k-k_0)^2/2a_0^2}, \quad (4.1.4)$$

which is also a Gaussian function. As shown in Figure 4.2, the Fourier transform has a single peak at wave number k_0 . The width of this peak, given by the parameter a_0 , is inversely proportional to the width of the wave packet. It is easy to show that the product of the mean-square width of the function, $\langle \Delta x^2 \rangle$, and the mean-square width of the Fourier transform, $\langle \Delta k^2 \rangle$, obeys the inequality

$$\langle \Delta x^2 \rangle \langle \Delta k^2 \rangle \geq \frac{1}{4}. \quad (4.1.5)$$

The above inequality is often referred to as the uncertainty theorem for Fourier transforms, and is analogous to the well-known uncertainty theorem of quantum mechanics [Townsend, 1992]. The Gaussian harmonic packet is unique because $\langle \Delta x^2 \rangle \langle \Delta k^2 \rangle$ is exactly equal to 1/4, which is the smallest allowed value.

The uncertainty theorem simply expresses the fact that a sharply peaked transform corresponds to a wave packet that is spread out over a broad region in space, and vice versa. An extreme example is obtained by taking the limit of equations (4.1.3) and (4.1.4) as the parameter a_0 goes to zero. In this case, the wave packet (4.1.3) becomes a plane wave,

$$f(x) = e^{ik_0 x}, \quad (4.1.6)$$

which has a constant amplitude extending from $-\infty$ to $+\infty$, and the Fourier transform (4.1.4) becomes a Dirac delta function

$$\tilde{f}(k) = \sqrt{2\pi} \delta(k - k_0), \quad (4.1.7)$$

which is sharply localized in k -space, being non-zero only when $k = k_0$.

By a simple four-dimensional extension of expression (4.1.1), we can represent any function of space and time, $f(\mathbf{r}, t)$, by the four-dimensional integral

$$f(\mathbf{r}, t) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \tilde{f}(\mathbf{k}, \omega) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} d^3k d\omega, \quad (4.1.8)$$

where $d^3k = dk_x dk_y dk_z$ and $\mathbf{k} \cdot \mathbf{r} = k_x x + k_y y + k_z z$. The quantity ω is the frequency and is related to the time coordinate t in the same way that the wave number components, k_x, k_y and k_z , are related to the spatial coordinates, x, y and z . The function $\exp[i(\mathbf{k} \cdot \mathbf{r} - \omega t)]$ in the integrand represents an elementary plane wave. As shown in Figure 4.3, the vector \mathbf{k} is perpendicular to the planes of constant phase ($\mathbf{k} \cdot \mathbf{r} - \omega t = \text{constant}$). The minus sign is placed in front of ω so that the planes of constant phase move in the \mathbf{k} direction. The vector \mathbf{k} is often called the wave vector or propagation vector, since it indicates the direction of motion of the planes of constant phase. The velocity with which the planes of constant phase move is called the phase velocity and is given by

$$\mathbf{v}_p = \frac{\omega}{k} \hat{\mathbf{k}}. \quad (4.1.9)$$

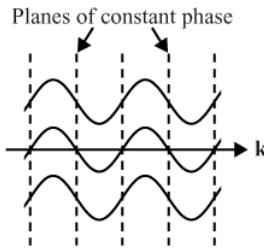


Fig. 4.3 An elementary plane wave with wave vector \mathbf{k} .

4.1.1 The dispersion relation

A wave $f(\mathbf{r}, t)$ propagating in some medium usually satisfies a differential equation that determines the possible solutions for $f(\mathbf{r}, t)$ in that medium. If the differential equation is linear, Fourier analysis provides a simple method of determining the allowed solutions. This simplification arises because the Fourier transform of the derivative $\partial f / \partial t$ can be obtained by simply multiplying the transform of f by $-i\omega$. This property can be demonstrated by differentiating both sides of equation (4.1.8) with respect to t , which gives

$$\frac{\partial f}{\partial t} = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} (-i\omega \tilde{f}) e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} d^3k d\omega. \quad (4.1.10)$$

The Fourier transform of $\partial f / \partial t$ is thus seen to be $(-i\omega \tilde{f})$. A similar derivation can be performed to show that the Fourier transform of ∇f is $i\mathbf{k}\tilde{f}$. Therefore, any combination of derivatives can be represented by the following operator substitutions

$$\begin{aligned}\frac{\partial}{\partial t} &\rightarrow -i\omega \\ \nabla &\rightarrow ik.\end{aligned}\tag{4.1.11}$$

It is readily shown that these operator substitutions hold for any sequence of operations, including all operations involving vector calculus. Thus, $\partial^n f / \partial t^n \rightarrow (-i\omega)^n \tilde{f}$, $\nabla^2 f \rightarrow (ik)^2 \tilde{f}$, $\nabla \cdot \mathbf{f} \rightarrow ik \cdot \tilde{\mathbf{f}}$, $\nabla \times \mathbf{f} \rightarrow ik \times \tilde{\mathbf{f}}$, etc. It follows then that the Fourier transform of the homogeneous linear differential equation

$$D(\nabla, \partial/\partial t)f = 0\tag{4.1.12}$$

is simply

$$D(ik, -i\omega)\tilde{f} = 0.\tag{4.1.13}$$

The net effect is to convert the original differential equation (4.1.12) into an algebraic equation (4.1.13) involving ω , \mathbf{k} , and \tilde{f} . The resulting algebraic equation is usually much easier to solve than the original differential equation. For the homogeneous equation (4.1.13) to have a non-trivial solution for \tilde{f} the quantity $D(ik, -i\omega)$ must be zero. This condition, written as $D(\mathbf{k}, \omega) = 0$, is called the dispersion relation. The dispersion relation provides a relationship between ω and \mathbf{k} . In many cases of physical interest, the dispersion relation has multiple, discrete roots for ω , represented by

$$\omega = \omega_\alpha(\mathbf{k}), \quad \alpha = 1, 2, \dots, N.\tag{4.1.14}$$

These roots are called the normal modes. Except for degenerate cases, the number of normal modes is equal to the order of the differential equation. For example, the simple wave equation

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0,\tag{4.1.15}$$

where c is the speed of light, can be Fourier transformed to yield

$$\left[-k^2 + \frac{\omega^2}{c^2} \right] \tilde{f} = 0,\tag{4.1.16}$$

which has a dispersion relation

$$D(\mathbf{k}, \omega) = -k^2 + \frac{\omega^2}{c^2} = 0,\tag{4.1.17}$$

with roots $\omega = \pm ck$. The two normal modes in this case represent waves traveling to the right and left with phase velocities $v_p = \pm c$.

Once the roots of the dispersion relation are known, the principle of superposition allows us to find the general solution to the linear differential equation (4.1.12) by integrating over \mathbf{k} and summing over all the normal modes. Thus, the general solution can be represented by

$$f(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}} \sum_{\alpha=1}^N \int_{-\infty}^{\infty} \tilde{f}_\alpha(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{r} - \omega_\alpha(\mathbf{k})t]} d^3k. \quad (4.1.18)$$

The roots $\omega_\alpha(\mathbf{k})$, specified by equation (4.1.14), determine the propagation speed of the individual plane waves that make up the wave packet $f(\mathbf{r}, t)$. Because the plane waves usually propagate at different speeds, the wave packet generally tends to spread out in space with increasing time. This spreading is called dispersion.

For a vector wave field \mathbf{f} , it is usually necessary to specify a system of linear homogeneous differential equations that can be written in the following matrix form

$$\overleftrightarrow{D}(\nabla, \partial/\partial t) \cdot \mathbf{f} = 0. \quad (4.1.19)$$

Using the operator substitution given by equation (4.1.11), the corresponding Fourier transform is then

$$\overleftrightarrow{D}(i\mathbf{k}, -i\omega) \cdot \tilde{\mathbf{f}} = 0 \quad (4.1.20)$$

It is well known that a homogeneous system of linear equations has a non-trivial solution if and only if the determinant of the matrix is zero, which gives the dispersion relation

$$D(\mathbf{k}, \omega) = \text{Det}[\overleftrightarrow{D}(i\mathbf{k}, -i\omega)] = 0. \quad (4.1.21)$$

Each root of the dispersion relation has an associated vector $\tilde{\mathbf{f}}$, that is called the eigenvector for that mode. The eigenvector characterizes the field geometry associated with that mode of propagation.

4.1.2 Group velocity

Usually the integrals involved in the Fourier representation of a wave are quite complicated, since the individual plane waves are propagating in different directions at different speeds. However, there is one particular situation for which the waveform has a simple representation. If the Fourier transform is sharply peaked around a particular wave number (such as the Gaussian harmonic wave packet with $a_0 \ll k_0$), the shape of the envelope of the wave packet is preserved to a first approximation. The velocity at which the envelope moves is called the group velocity.

To obtain an expression for the group velocity, we note that if the Fourier transform is sharply peaked at $\mathbf{k} = \mathbf{k}_0$ for one of the modes, only waves near \mathbf{k}_0 contribute significantly to the wave packet for that mode. We can then obtain a good approximation to the integral in equation (4.1.8) by performing a Taylor series expansion of the dispersion relation around k_0 and keeping only the first two terms

$$\omega(\mathbf{k}) = \omega_0 + \nabla_{\mathbf{k}}\omega \cdot (\mathbf{k} - \mathbf{k}_0) + \dots \quad (4.1.22)$$

where ω_0 and $\nabla_{\mathbf{k}}\omega = \hat{\mathbf{x}} \partial\omega/\partial k_x + \hat{\mathbf{y}} \partial\omega/\partial k_y + \hat{\mathbf{z}} \partial\omega/\partial k_z$, are evaluated at $\mathbf{k} = \mathbf{k}_0$. As shown in Figure 4.4, the first two terms represent a straight line approximation to the dispersion

relation in the vicinity of the peak in $\tilde{f}(\mathbf{k})$. Substituting equation (4.1.22) into the integral representation (4.1.18) and neglecting the higher order terms, we obtain

$$f(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \tilde{f}(\mathbf{k}) e^{i[\mathbf{k} \cdot \mathbf{r} - \omega_0 t - \nabla_{\mathbf{k}} \omega \cdot (\mathbf{k} - \mathbf{k}_0)t]} d^3 k. \quad (4.1.23)$$

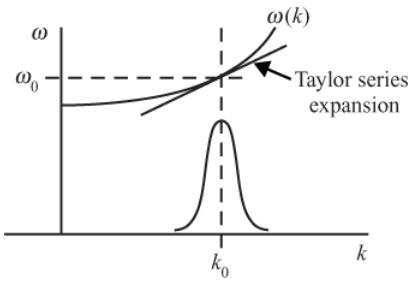


Fig. 4.4 Taylor series expansion of $\omega(k)$ around the point k_0, ω_0 .

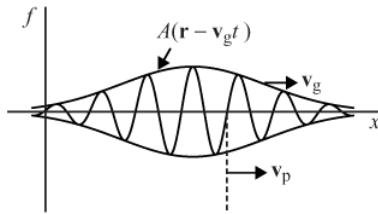


Fig. 4.5 The envelope of a wave packet moves at the group velocity and the planes of constant phase move at the phase velocity.

After rearranging terms, the integral on the right-hand side of the equation can be rewritten

$$f(\mathbf{r}, t) = \left[\frac{1}{(2\pi)^{3/2}} \int_{-\infty}^{\infty} \tilde{f}(\mathbf{k}) e^{i(\mathbf{k} - \mathbf{k}_0) \cdot (\mathbf{r} - \mathbf{v}_g t)} d^3 k \right] e^{i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)}, \quad (4.1.24)$$

where $\mathbf{v}_g = \nabla_{\mathbf{k}} \omega$ is called the group velocity. It is easy to see that the above equation can be written in the form

$$f(\mathbf{r}, t) = A(\mathbf{r} - \mathbf{v}_g t) e^{i(\mathbf{k}_0 \cdot \mathbf{r} - \omega_0 t)}. \quad (4.1.25)$$

This equation shows that the waveform for a given mode consists of a plane wave of wave vector \mathbf{k}_0 and frequency ω_0 modulated by an amplitude factor $A(\mathbf{r} - \mathbf{v}_g t)$. A typical waveform is shown in Figure 4.5. The wave packet consists of a plane wave that moves at the phase velocity $\mathbf{v}_p = (\omega_0/k_0)\hat{\mathbf{k}}$, and an outer envelope that moves at the group velocity $\mathbf{v}_g = \nabla_{\mathbf{k}} \omega$. Since energy is carried by the wave packet, the group velocity is the velocity of energy propagation.

Several interesting features can be noted about the group velocity. Since the group velocity and phase velocity are in general different, the planes of constant phase move through the wave packet. A given wave crest starts with a small amplitude at one end of the packet,

grows in amplitude as it reaches the center of the packet, and finally disappears at the other end. An exception to this behavior occurs for waves that obey the simple wave equation

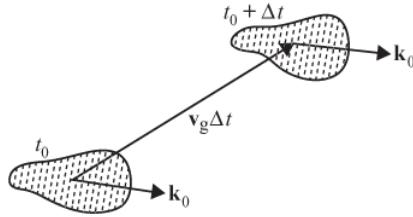


Fig. 4.6 In an anisotropic medium the group velocity, v_g , may be in a direction totally different than the wave vector, \mathbf{k}_0 .

$$\frac{\partial^2 f}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0, \quad (4.1.26)$$

for which the phase and group velocities are the same, i.e., $v_p = v_g = c$. In this special case, the shape of the waveform stays constant in time. Such waves are said to be non-dispersive, since the wave packet does not tend to spread out with increasing time. Electromagnetic waves in free space and sound waves in air have this property.

If the medium is anisotropic the group velocity and phase velocity may not even be in the same direction. In such an anisotropic medium the wave packet moves at an angle relative to the wave vector, as shown in Figure 4.6. This unusual behavior occurs in doubly refracting optical crystals and in plasmas with a static magnetic field. In order to follow the direction of wave energy flow in an inhomogeneous medium it is useful to introduce the concept of a ray path. The ray path is a sequence of connected line elements that are everywhere in the direction of the local group velocity. To construct the ray path it is useful to have a simple method of visualizing the direction of the group velocity. The group velocity direction can be visualized by noting that \mathbf{v}_g is simply the gradient of $\omega(\mathbf{k})$ in \mathbf{k} space

$$\mathbf{v}_g = \nabla_{\mathbf{k}} \omega(\mathbf{k}) = \hat{x} \frac{\partial \omega}{\partial k_x} + \hat{y} \frac{\partial \omega}{\partial k_y} + \hat{z} \frac{\partial \omega}{\partial k_z}. \quad (4.1.27)$$

From the properties of the gradient operator it follows that the group velocity is perpendicular to a contour of constant ω in \mathbf{k} -space. This relationship is illustrated by the geometric construction in Figure 4.7, which shows that the group velocity is perpendicular to a contour of constant ω . In practice, it is often more convenient to deal with a dimensionless index of refraction vector, $\mathbf{n} = c\mathbf{k}/\omega$, rather than the wave vector \mathbf{k} . Let us assume that the index of refraction is given as a function of the angle θ , called the wave normal angle, between the wave vector and some preferred axis of symmetry, typically the magnetic field, \mathbf{B}_0 . The index of refraction for constant ω can then be represented as a curve $\mathbf{n}(\theta)$ in polar form, as shown in Figure 4.8. The surface generated by rotating $\mathbf{n}(\theta)$ around the symmetry axis \mathbf{B}_0 is called the index of refraction surface. Since ω is constant along this surface, it follows from our previous discussion that the group velocity is perpendicular to

the index of refraction surface, i.e., in the direction of the gradient $\nabla_{\mathbf{n}}\omega(\mathbf{n})$. Using $\hat{\mathbf{k}}$ and $\hat{\theta}$ unit vectors, it is easy to show that the group velocity can be written in terms of this gradient as

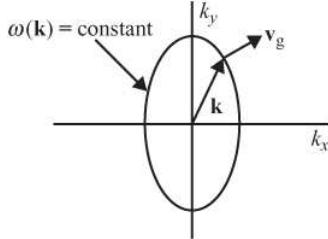


Fig. 4.7 In \mathbf{k} -space, the group velocity is perpendicular to a surface of constant ω .

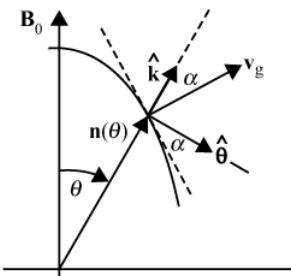


Fig. 4.8 Geometric construction used to compute the angle α between the group velocity and the wave vector.

$$\mathbf{v}_g = \frac{c}{\omega} \nabla_{\mathbf{n}}\omega(\mathbf{n}) = \frac{c}{\omega} \left[\hat{\mathbf{k}} \frac{\partial \omega}{\partial n} + \hat{\theta} \frac{1}{n} \frac{\partial \omega}{\partial \theta} \right]. \quad (4.1.28)$$

The angle α between the group velocity and the wave vector is then given by

$$\tan \alpha = \frac{\frac{1}{n} \left(\frac{\partial \omega}{\partial \theta} \right)_n}{\left(\frac{\partial \omega}{\partial n} \right)_\theta}, \quad (4.1.29)$$

where the subscripts n and θ indicate that the derivatives are to be computed with n and θ held constant. Using the following theorem relating the partial derivatives of three implicit functions ω , n , and θ [Kaplan, 1952]

$$\left(\frac{\partial \omega}{\partial \theta} \right)_n \left(\frac{\partial \theta}{\partial n} \right)_\omega \left(\frac{\partial n}{\partial \omega} \right)_\theta = -1, \quad (4.1.30)$$

equation (4.1.29) simplifies to

$$\tan \alpha = -\frac{1}{n} \left(\frac{\partial n}{\partial \theta} \right)_\omega. \quad (4.1.31)$$

This relationship provides a convenient method of calculating the group velocity direction if the index of refraction is known as a function of θ .

4.2 General form of the dispersion relation

To analyze the waves that can exist in a plasma, the basic equations that must be solved self-consistently are Maxwell's equations and the particle equations of motion. As is well known, Maxwell's equations can be written in two equivalent forms, the so-called "microscopic" and "macroscopic" forms. These two forms are summarized below:

Microscopic	Macroscopic	
$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t}$	$\nabla \times \mathbf{H} = \mathbf{J}_r + \frac{\partial \mathbf{D}}{\partial t}$	
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	
$\nabla \cdot \mathbf{E} = \frac{\rho_q}{\epsilon_0}$	$\nabla \cdot \mathbf{D} = \rho_r$	
$\nabla \cdot \mathbf{B} = 0$	$\nabla \cdot \mathbf{B} = 0$	(4.2.1)

In the macroscopic approach the charge is divided into a real charge and a polarization charge, $\rho_q = \rho_r + \rho_p$, and the current is divided into a real current and a magnetization current, $\mathbf{J} = \mathbf{J}_r + \mathbf{J}_m$. This division is completely arbitrary, depending on how we define real charges and polarization charges, and real currents and magnetization currents. For our purposes it is useful to count all of the charges in the plasma as polarization charges, so that the displacement field is given by $\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}$, where \mathbf{P} is the dipole moment per unit volume. All of the currents in the plasma are then included in the displacement current $\partial \mathbf{D} / \partial t$. Since the magnetic moment of the individual particles is normally negligible, in a plasma we usually assume that $\mathbf{B} = \mu_0 \mathbf{H}$.

4.2.1 The conductivity and dielectric tensors

To obtain a closed system of equations using the macroscopic approach, it is necessary to specify the relationship between \mathbf{D} and \mathbf{E} . This specification must ultimately come from the equations of motion of the particles. The procedure is first to solve the equations of motion for some assumed wave field. Since the equations of motion are assumed to be linear, this solution is most easily obtained by using Fourier transforms, which gives the Fourier transform of the velocity, $\tilde{\mathbf{v}}$, as a linear function of the Fourier transform of the electric field, $\tilde{\mathbf{E}}$. The solution for $\tilde{\mathbf{v}}$ can then be substituted into the equation for the current density

$$\tilde{\mathbf{J}} = \sum_s n_s e_s \tilde{\mathbf{v}}_s, \quad (4.2.2)$$

and organized into an equation of the form

$$\tilde{\mathbf{J}} = \vec{\sigma} \cdot \tilde{\mathbf{E}}, \quad (4.2.3)$$

where the matrix $\vec{\sigma}$ is a tensor called the conductivity tensor.

Once the conductivity tensor has been determined, the next step is to relate the electric

field $\tilde{\mathbf{E}}$ to the displacement vector $\tilde{\mathbf{D}}$. Since the system is assumed to be linear, in the macroscopic approach the displacement vector $\tilde{\mathbf{D}}$ must be a linear function of the electric field $\tilde{\mathbf{E}}$. Since we anticipate that anisotropic effects are likely to occur, the linear relationship between $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{E}}$ is assumed to be of the form

$$\tilde{\mathbf{D}} = \epsilon_0 \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}}, \quad (4.2.4)$$

where the matrix $\overset{\leftrightarrow}{\mathbf{K}}$ is called the dielectric tensor. The above equation represents a generalization of the relation $\mathbf{D} = \epsilon_0 K \mathbf{E}$, commonly used in the description of isotropic media.

The relationship between the dielectric tensor and the conductivity tensor can be obtained by comparing the microscopic and macroscopic forms of Ampère's law, which in Fourier transformation form can be written

Microscopic	Macroscopic
$i\mathbf{k} \times \tilde{\mathbf{B}} = \mu_0 \overset{\leftrightarrow}{\sigma} \cdot \tilde{\mathbf{E}} + \epsilon_0 \mu_0 (-i\omega) \tilde{\mathbf{E}}$,	$i\mathbf{k} \times \tilde{\mathbf{B}} = \epsilon_0 \mu_0 (-i\omega) \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}}$.

(4.2.5)

Equating the right-hand sides of the above two equations gives the following relationship between the dielectric tensor and the conductivity tensor:

$$\overset{\leftrightarrow}{\mathbf{K}} = \overset{\leftrightarrow}{\mathbf{1}} - \frac{\overset{\leftrightarrow}{\sigma}}{i\omega\epsilon_0}, \quad (4.2.6)$$

where $\overset{\leftrightarrow}{\mathbf{1}}$ is the unit tensor.

The procedure for computing the dielectric tensor is first to compute the conductivity tensor using the particle equation of motion and equation (4.2.2) for the current density. The dielectric tensor is then computed using equation (4.2.6). This procedure is quite general, and works for any plasma, hot or cold, provided only that the plasma can be described by a linear system of equations.

4.2.2 The homogeneous equation

To obtain the dispersion relation, we need to look for a solution to Maxwell's equations, ignoring external sources. A homogeneous (source free) system of equations can be obtained by eliminating either \mathbf{E} or \mathbf{B} from Faraday's law and Ampere's law (4.2.1). After Fourier transforming, Faraday's and Ampère's laws become (using $\mu_0\epsilon_0 = 1/c^2$)

$$i\mathbf{k} \times \tilde{\mathbf{E}} = -(-i\omega)\tilde{\mathbf{B}} \quad \text{and} \quad i\mathbf{k} \times \tilde{\mathbf{B}} = \frac{-i\omega}{c^2} \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}}. \quad (4.2.7)$$

Eliminating $\tilde{\mathbf{B}}$ between these two equations gives a homogeneous equation for the electric field

$$\mathbf{k} \times (\mathbf{k} \times \tilde{\mathbf{E}}) + \frac{\omega^2}{c^2} \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}} = 0. \quad (4.2.8)$$

Using the definition of the index of refraction, $\mathbf{n} = c\mathbf{k}/\omega$, the above equation can be expressed in the following, somewhat simpler, form

$$\mathbf{n} \times (\mathbf{n} \times \tilde{\mathbf{E}}) + \overleftrightarrow{\mathbf{K}} \cdot \tilde{\mathbf{E}} = 0. \quad (4.2.9)$$

Either of the above two homogeneous equations ((4.2.8) or (4.2.9)) can be written in matrix form as $\overset{\leftrightarrow}{D} \cdot \tilde{\mathbf{E}} = 0$. A non-trivial solution for $\tilde{\mathbf{E}}$ is then possible if and only if the determinant of the matrix is zero, which gives the dispersion relation, $D(\mathbf{k}, \omega)$. The electric field eigenvector associated with each root of the dispersion relation can be obtained from the homogeneous equation ((4.2.8) or (4.2.9)), and the corresponding magnetic field eigenvector can be obtained from Faraday's law, $\tilde{\mathbf{B}} = (\mathbf{k}/\omega) \times \tilde{\mathbf{E}}$ or $c\tilde{\mathbf{B}} = \mathbf{n} \times \tilde{\mathbf{E}}$.

4.3 Waves in a cold uniform unmagnetized plasma

To illustrate the above procedure we first consider the special case of waves propagating in a cold uniform unmagnetized plasma. To obtain a linear set of equations, all of the dependent variables (\mathbf{v}_s , n_s , \mathbf{E} and \mathbf{B}) are assumed to consist of a constant uniform zero-order term plus a small first-order perturbation. For example, the number density is assumed to be a constant uniform density plus a small perturbation, $n_s = n_{s0} + n_{s1}$. Since the plasma is cold (i.e., zero temperature), the zero-order velocities are assumed to be zero. The zero-order electric field must also be zero, otherwise at zero order the particles would not remain at rest. Since the plasma is unmagnetized, the zero-order magnetic field is also assumed to be zero. Based on these assumptions we can then write

$$\begin{aligned} n_s &= n_{s0} + n_{s1} \\ \mathbf{v}_s &= \mathbf{v}_{s1} \\ \mathbf{E} &= \mathbf{E}_1 \\ \mathbf{B} &= \mathbf{B}_1. \end{aligned} \quad (4.3.1)$$

In these equations the subscript (0) identifies the zero-order quantities and the subscript (1) identifies a first-order quantity. To assure charge neutrality for the unperturbed system, we must also require that $\sum_s e_s n_{s0} = 0$.

Next we solve the equation of motion and compute the current for an assumed first-order wave field. The equation of motion for a particle of mass m_s and charge e_s is

$$m_s \frac{d\mathbf{v}_{s1}}{dt} = e_s [\mathbf{E}_1 + \mathbf{v}_{s1} \times \mathbf{B}_1] \quad (4.3.2)$$

and the equation for the current density is

$$\mathbf{J} = \sum_s n_{s0} e_s \mathbf{v}_{s1} + \sum_s n_{s1} e_s \mathbf{v}_{s1}. \quad (4.3.3)$$

In these equations the non-linear terms $\mathbf{v}_{s1} \times \mathbf{B}_1$ and $n_{s1} \mathbf{v}_{s1}$ are of second order because

they represent the product of two first-order terms. If the amplitude of the first-order perturbation is assumed to be small, so that $|\mathbf{v}_{s1} \times \mathbf{B}_1| \ll |\mathbf{E}_1|$ and $|\mathbf{n}_{s1}\mathbf{v}_{s1}| \ll |\mathbf{n}_{s0}\mathbf{v}_{s1}|$, then the second-order terms can be neglected compared with the first-order terms. The term $d\mathbf{v}_{s1}/dt$ represents the acceleration evaluated in the particle frame of reference. Because the particles are initially at rest, $d\mathbf{v}_{s1}/dt$ can be replaced by $\partial\mathbf{v}_{s1}/\partial t$. The linearized equations then become

$$m_s \frac{\partial \mathbf{v}_s}{\partial t} = e_s \mathbf{E}, \quad (4.3.4)$$

and

$$\mathbf{J} = \sum_s n_{s0} e_s \mathbf{v}_s, \quad (4.3.5)$$

where for notational simplicity we have dropped the subscript 1 on the first-order terms. The subscript 0 on the zero-order term can always be used to distinguish the zero-order terms from the first-order terms.

Next we compute the conductivity tensor. The Fourier transforms of the equation of motion (4.3.4) and the current density (4.3.5) are

$$m_s(-i\omega)\tilde{\mathbf{v}}_s = e_s \tilde{\mathbf{E}} \quad (4.3.6)$$

and

$$\tilde{\mathbf{J}} = \sum_s n_{s0} e_s \tilde{\mathbf{v}}_s. \quad (4.3.7)$$

Solving equation (4.3.6) for the velocity and substituting into the right-hand side of equation (4.3.7) gives

$$\tilde{\mathbf{J}} = \sum_s \frac{n_{s0} e_s^2}{(-i\omega)m_s} \tilde{\mathbf{E}}. \quad (4.3.8)$$

The conductivity tensor, defined by equation (4.2.3), can now be identified, and is given by

$$\vec{\sigma} = \vec{1} \sum_s \frac{n_{s0} e_s^2}{(-i\omega)m_s}. \quad (4.3.9)$$

In this case, the conductivity tensor is diagonal, which indicates an isotropic scalar conductivity. The conductivity tensor is also purely imaginary, which indicates that \mathbf{J} and \mathbf{E} are shifted in phase by $\pi/2$. This phase shift arises from the $\pi/2$ phase shift between the velocity and the electric field implied by the “i” in equation (4.3.6). The conductivity also decreases with increasing frequency. This frequency dependence occurs because at higher frequencies the particles have a shorter time to respond to the electric field, thereby causing smaller velocities and currents. Because of the very small mass of the electron, the dominant contribution to the conductivity tensor comes from the electrons.

Having determined the conductivity tensor, the dielectric tensor can be computed using equation (4.2.6) which gives

$$\vec{\mathbf{K}} = \vec{1} \left(1 - \frac{\omega_p^2}{\omega^2} \right), \quad (4.3.10)$$

where $\omega_p^2 = \sum_s \omega_{ps}^2$, and $\omega_{ps}^2 = n_{s0}e^2/\epsilon_0 m_s$. Note that since the electron mass is much smaller than the ion mass, the sum over species is dominated by the electron term, so to a good approximation $\omega_p^2 \approx \omega_{pe}^2$. Thus, electrons play the dominant role in determining the dielectric properties of the plasma.

To find the dispersion relation and the electric field eigenvectors, we must next analyze the homogeneous equation for the electric field. Substituting the dielectric tensor (4.3.10) into the homogeneous equation (4.2.8), we obtain the equation

$$c^2 \mathbf{k} \times (\mathbf{k} \times \tilde{\mathbf{E}}) + (\omega^2 - \omega_p^2) \tilde{\mathbf{E}} = 0. \quad (4.3.11)$$

Since the plasma is isotropic, without loss of generality we can choose a coordinate system with the z axis along the \mathbf{k} vector, so that $\mathbf{k} = (0, 0, k)$ and $\tilde{\mathbf{E}} = (\tilde{E}_x, \tilde{E}_y, \tilde{E}_z)$. Equation (4.3.11) can then be rewritten in the form

$$\begin{bmatrix} -c^2k^2 + \omega^2 - \omega_p^2 & 0 & 0 \\ 0 & -c^2k^2 + \omega^2 - \omega_p^2 & 0 \\ 0 & 0 & \omega^2 - \omega_p^2 \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (4.3.12)$$

The dispersion relation is obtained by setting the determinant of the matrix to zero, which gives

$$D(k, \omega) = (-c^2k^2 + \omega^2 - \omega_p^2)^2(\omega^2 - \omega_p^2) = 0. \quad (4.3.13)$$

The roots of the dispersion relation and the corresponding eigenvectors are easily shown to be

$$\omega^2 = \omega_p^2 + c^2k^2, \quad \tilde{\mathbf{E}} = (\tilde{E}_x, \tilde{E}_y, 0), \quad (4.3.14)$$

and

$$\omega^2 = \omega_p^2, \quad \tilde{\mathbf{E}} = (0, 0, \tilde{E}_z). \quad (4.3.15)$$

Since \mathbf{k} is in the z direction, the electric fields associated with the above two roots are, respectively, perpendicular and parallel to the direction of propagation. These roots are called the transverse ($\mathbf{k} \cdot \tilde{\mathbf{E}} = 0$) and longitudinal ($\mathbf{k} \times \tilde{\mathbf{E}} = 0$) modes.

4.3.1 The transverse mode

The electric field eigenvector for the transverse mode is $\tilde{\mathbf{E}} = (\tilde{E}_x, \tilde{E}_y, 0)$. As can be seen, the electric field eigenvector has two independent components, \tilde{E}_x , and \tilde{E}_y . Each of these electric field components has an associated magnetic field component given by $\tilde{\mathbf{B}} = (\mathbf{k}/\omega) \times \tilde{\mathbf{E}}$. For the \tilde{E}_x component, the \mathbf{k} vector, the electric field $\tilde{\mathbf{E}}$, and the magnetic field $\tilde{\mathbf{B}}$ form an orthogonal triad as shown in Figure 4.9. A similar triad exists for the \tilde{E}_y component, but rotated by $\pi/2$ with respect to the z axis.

Since the field geometry is such that $\mathbf{k} \cdot \tilde{\mathbf{E}} = 0$, Gauss' law, $\tilde{\rho}_q = \epsilon_0(\mathbf{i}\mathbf{k}) \cdot \tilde{\mathbf{E}}$, shows that there is no charge density fluctuation associated with the transverse mode. Thus, even

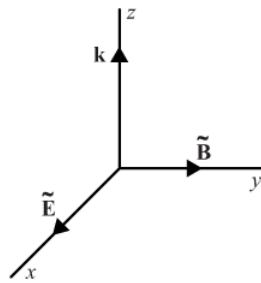


Fig. 4.9 The electric and magnetic field eigenvectors for the transverse electromagnetic mode.

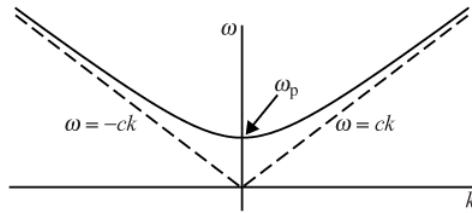


Fig. 4.10 The dispersion relation for the transverse electromagnetic mode.

though this mode generates an electrical current, the current does not cause a charge density fluctuation.

The dispersion relation for the transverse mode, $\omega^2 = \omega_p^2 + c^2 k^2$, is plotted in Figure 4.10. As can be seen, for large ω and k the dispersion relation is asymptotic to the straight lines $\omega = \pm ck$, which are the free space branches of the electromagnetic dispersion relation. Thus, at high frequencies the transverse mode reduces to the free space electromagnetic mode. No real solution for ω exists at frequencies below the plasma frequency ω_p . Therefore, propagation cannot occur for frequencies below the plasma frequency. The existence of this propagation cutoff can be easily demonstrated by solving the dispersion relation for the wave number

$$k = \pm \frac{1}{c} \sqrt{\omega^2 - \omega_p^2}. \quad (4.3.16)$$

For $|\omega| > \omega_p$ the wave number is real, which corresponds to a propagating wave. For $|\omega| < \omega_p$ the wave number is purely imaginary, $k = ik_i$, which corresponds to a pure exponential decay with no spatial oscillations. Such a wave is said to be *evanescent* or non-propagating. The reason for the cutoff at the plasma frequency can be traced to the relationship between the conduction current and the displacement current. It is easy to show that the ratio of the conduction current to the displacement current is given by

$$\frac{\text{conduction current}}{\text{displacement current}} = \frac{\sum_s \frac{n_{s0} e^2}{(-i\omega)m_s}}{\epsilon_0(-i\omega)} = -\frac{\omega_p^2}{\omega^2}. \quad (4.3.17)$$

The minus sign in the above equation indicates that the conduction current opposes the displacement current. At high frequencies, $\omega \gg \omega_p$, the conduction current has little effect

on the free space electromagnetic mode. However, because of the $1/\omega^2$ dependence, as the frequency decreases the conduction current increases, until at the plasma frequency it just cancels the displacement current. At and below this frequency propagation cannot occur, since the current on the right-hand side of Ampère's law (4.2.5) has the wrong sign.

The phase and group velocities can be computed from the dispersion relation and are given by

$$v_p = \frac{\omega}{k} = \frac{c}{\sqrt{1 - \frac{\omega_p^2}{\omega^2}}} \quad \text{and} \quad v_g = \frac{\partial \omega}{\partial k} = c \sqrt{1 - \frac{\omega_p^2}{\omega^2}}. \quad (4.3.18)$$

As shown in Figure 4.11, the phase velocity goes to infinity at the plasma frequency, and the group velocity goes to zero. Note that the group velocity, which is the velocity of energy propagation, is always less than the speed of light. It can also be verified that the group velocity and phase velocity satisfy the relationship $v_p v_g = c^2$. The phase and group velocities are also the same as the phase and group velocities for an electromagnetic wave propagating in a waveguide, with the plasma frequency playing the same role as the cutoff frequency of a waveguide.

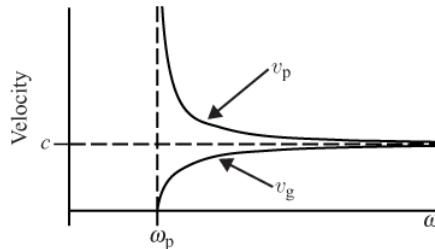


Fig. 4.11 The phase velocity (top panel) and group velocity (bottom panel) of the transverse electromagnetic mode.

4.3.2 The longitudinal mode

For the longitudinal mode, equation (4.3.15) shows that the dispersion relation is $\omega = \pm \omega_p$, independent of k . This mode corresponds to the electron plasma oscillations analyzed in Section ?? using a simple slab model. The wave magnetic field of this mode is zero because $\tilde{\mathbf{B}} = (\mathbf{k}/\omega) \times \tilde{\mathbf{E}} = 0$. However, the perturbed charge density is not zero, since $\tilde{\rho}_q = \epsilon_0(\mathbf{i}\mathbf{k}) \cdot \tilde{\mathbf{E}} \neq 0$. Thus, the longitudinal mode is associated with a charge density fluctuation. Since $\mathbf{k} \times \tilde{\mathbf{E}} = 0$ (or $\nabla \times \mathbf{E} = 0$), it also follows that the electric field of the longitudinal mode can be written as the gradient of an electrostatic potential, $\mathbf{E} = -\nabla\Phi$. Thus, the oscillations are purely electrostatic. It is also easy to see that the group velocity, $\partial\omega/\partial k$, is zero, since the oscillation frequency is independent of the wave number. Thus, the envelope of the wave packet does not propagate. Since the wave magnetic field is zero, the Poynting flux, $\mathbf{S} = (1/\mu_0) \mathbf{E} \times \mathbf{B}$, which represents the flow of electromagnetic energy, is also zero, consistent with the fact that the group velocity is zero. In the next chapter,

when non-zero temperatures are considered, we will show that the longitudinal mode has a non-zero group velocity.

4.3.3 External sources

The discussions in the previous two sections have been mainly concerned with the wave modes that can exist in a plasma with no external source. An external source can be easily incorporated by adding an external current density, \mathbf{J}_{ext} , to Ampère's law so that

$$\nabla \times \mathbf{B} = \mu_0 \frac{\partial \mathbf{D}}{\partial t} + \mu_0 \mathbf{J}_{\text{ext}}. \quad (4.3.19)$$

After Fourier transforming, eliminating $\tilde{\mathbf{B}}$ using Faraday's law, and introducing the dielectric tensor via $\tilde{\mathbf{D}} = \epsilon_0 \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}}$, the above equation can be written

$$\mathbf{k} \times (\mathbf{k} \times \tilde{\mathbf{E}}) + \frac{\omega^2}{c^2} \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}} = -i\omega \mu_0 \tilde{\mathbf{J}}_{\text{ext}}, \quad (4.3.20)$$

where $\tilde{\mathbf{J}}_{\text{ext}}$ is the Fourier transform of the external current distribution. Splitting the electric field and current density into longitudinal (parallel to \mathbf{k}) and transverse (perpendicular to \mathbf{k}) components, $\tilde{\mathbf{E}} = \tilde{\mathbf{E}}_{\parallel} + \tilde{\mathbf{E}}_{\perp}$, and $\tilde{\mathbf{J}} = \tilde{\mathbf{J}}_{\parallel} + \tilde{\mathbf{J}}_{\perp}$, and using the dielectric tensor given by equation (4.3.10), it is easy to show that the above equation separates nicely into two equations

$$(-c^2 k^2 + \omega^2 - \omega_p^2)^2 \tilde{\mathbf{E}}_{\perp} = -\frac{i\omega}{\epsilon_0} \tilde{\mathbf{J}}_{\perp \text{ext}} \quad (4.3.21)$$

and

$$(\omega^2 - \omega_p^2)^2 \tilde{\mathbf{E}}_{\parallel} = -\frac{i\omega}{\epsilon_0} \tilde{\mathbf{J}}_{\parallel \text{ext}}. \quad (4.3.22)$$

These two equations show that the transverse component of the external current acts as a source for the transverse electromagnetic waves, and the longitudinal component acts as a source for the longitudinal electrostatic waves.

In practice, it is somewhat difficult to interpret equations (4.3.21) and (4.3.22), since $\tilde{\mathbf{J}}_{\perp}$ and $\tilde{\mathbf{J}}_{\parallel}$ involve Fourier transforms. The interpretation is simplified somewhat by noting, from the continuity equation $\partial \rho_q / \partial t + \nabla \cdot \mathbf{J} = 0$, which after Fourier transforming becomes $-\omega \tilde{\rho}_q + \mathbf{k} \cdot \tilde{\mathbf{J}}_{\parallel} = 0$, that $\tilde{\mathbf{J}}_{\parallel}$ is associated with a charge density fluctuation, whereas $\tilde{\mathbf{J}}_{\perp}$ is not. Therefore, longitudinal waves can only be excited if charge density fluctuations are present in the source. For example, a loop antenna, such as shown in Figure 4.12, can only excite transverse waves, since there are no charge density fluctuations associated with this current system. On the other hand, an electric dipole antenna, such as shown in Figure 4.13, has both a transverse current and a longitudinal current (i.e., charge fluctuation) and can excite both transverse and longitudinal waves.

Once the Fourier transform of the source current is known, equations (4.3.21) and (4.3.22) can be solved for the electric fields by computing the inverse Fourier transforms of $\tilde{\mathbf{E}}_{\perp}$ and $\tilde{\mathbf{E}}_{\parallel}$. Note that the terms $-c^2 k^2 + \omega^2 - \omega_p^2$ and $\omega^2 - \omega_p^2$ occur in the denominators of the inverse

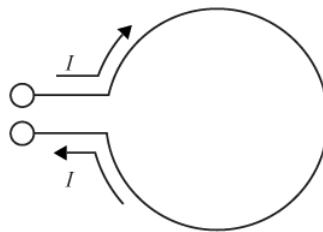


Fig. 4.12 A loop antenna has no charge density fluctuations and cannot excite the longitudinal electrostatic mode.

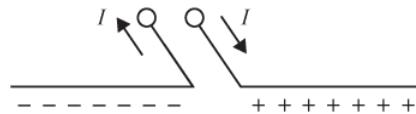


Fig. 4.13 An electric dipole antenna has a charge density fluctuation and can excite both longitudinal and transverse modes.

Fourier transforms. These denominators play a crucial role in determining the properties of the inverse Fourier transform by introducing poles in the integrals. The procedures for computing integrals of this type will be discussed further in Chapter ??.

4.4 Waves in a cold uniform magnetized plasma

We next consider the case of a cold plasma with an externally imposed static uniform magnetic field. Using the same linearization procedure as in Section 4.3, the magnetic field must now be written $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1$, where \mathbf{B}_0 is the externally imposed magnetic field and \mathbf{B}_1 is the first-order perturbation. The linearized equation of motion is then

$$m_s \frac{\partial \mathbf{v}_{s1}}{\partial t} = e_s [\mathbf{E}_1 + \mathbf{v}_{s1} \times \mathbf{B}_0], \quad (4.4.1)$$

where the term $\mathbf{v}_{s1} \times \mathbf{B}_1$ has been dropped because it is a product of two small first-order terms. Without loss of generality we can let the static magnetic field be in the z direction, $\mathbf{B}_0 = (0, 0, B_0)$. After Fourier transforming, the above equation becomes

$$\begin{aligned} -i\omega m_s \tilde{v}_{sx} &= e_s [\tilde{E}_x + \tilde{v}_{sy} B_0], \\ -i\omega m_s \tilde{v}_{sy} &= e_s [\tilde{E}_y + \tilde{v}_{sx} B_0], \\ -i\omega m_s \tilde{v}_{sz} &= e_s \tilde{E}_z, \end{aligned} \quad (4.4.2)$$

where for notational simplicity the subscript 1 on the first-order terms has again been dropped. By introducing the cyclotron frequency, $\omega_{cs} = e_s B_0 / m_s$, the above equations can be written in matrix form as

$$\begin{bmatrix} -i\omega & -\omega_{cs} & 0 \\ -\omega_{cs} & -i\omega & 0 \\ 0 & 0 & -i\omega \end{bmatrix} \begin{bmatrix} \tilde{v}_{sx} \\ \tilde{v}_{sy} \\ \tilde{v}_{sz} \end{bmatrix} = \frac{e_s}{m_s} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}. \quad (4.4.3)$$

This set of linear equations can be solved for the velocity components \tilde{v}_{sx} , \tilde{v}_{sy} , and \tilde{v}_{sz} by inverting the matrix on the left-hand side, which gives

$$\begin{bmatrix} \tilde{v}_{sx} \\ \tilde{v}_{sy} \\ \tilde{v}_{sz} \end{bmatrix} = \frac{e_s}{m_s} \begin{bmatrix} \frac{-i\omega}{\omega_{cs}^2 - \omega^2} & \frac{\omega_{cs}}{\omega_{cs}^2 - \omega^2} & 0 \\ \frac{-\omega_{cs}}{\omega_{cs}^2 - \omega^2} & \frac{-i\omega}{\omega_{cs}^2 - \omega^2} & 0 \\ 0 & 0 & \frac{i}{\omega} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}. \quad (4.4.4)$$

Following the same procedure as in the previous section, the conductivity tensor can be determined by computing the current, $\tilde{\mathbf{J}} = \sum_s n_{s0} e_s \tilde{\mathbf{v}}$, which in matrix form becomes

$$\begin{bmatrix} \tilde{J}_x \\ \tilde{J}_y \\ \tilde{J}_z \end{bmatrix} = \sum_s \frac{n_{s0} e_s^2}{m_s} \begin{bmatrix} \frac{-i\omega}{\omega_{cs}^2 - \omega^2} & \frac{\omega_{cs}}{\omega_{cs}^2 - \omega^2} & 0 \\ \frac{-\omega_{cs}}{\omega_{cs}^2 - \omega^2} & \frac{-i\omega}{\omega_{cs}^2 - \omega^2} & 0 \\ 0 & 0 & \frac{i}{\omega} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}. \quad (4.4.5)$$

The conductivity tensor, defined by equation (4.2.3), can now be identified and is given by

$$\sigma = \sum_s \frac{n_{s0} e_s^2}{m_s} \begin{bmatrix} \frac{-i\omega}{\omega_{cs}^2 - \omega^2} & \frac{\omega_{cs}}{\omega_{cs}^2 - \omega^2} & 0 \\ \frac{-\omega_{cs}}{\omega_{cs}^2 - \omega^2} & \frac{-i\omega}{\omega_{cs}^2 - \omega^2} & 0 \\ 0 & 0 & \frac{i}{\omega} \end{bmatrix}. \quad (4.4.6)$$

Finally, the dielectric tensor, defined by equation (4.2.6), can be computed and has the form

$$\overset{\leftrightarrow}{\mathbf{K}} = \begin{bmatrix} S & -iD & 0 \\ -iD & S & 0 \\ 0 & 0 & P \end{bmatrix}, \quad (4.4.7)$$

where

$$S = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2 - \omega_{cs}^2}, \quad D = \sum_s \frac{\omega_{cs} \omega_{ps}^2}{\omega(\omega^2 - \omega_{cs}^2)} \quad (4.4.8)$$

and

$$P = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2}. \quad (4.4.9)$$

The term P is seen to be identical to the dielectric tensor elements for an unmagnetized plasma, see equation (4.3.10). This similarity occurs because for motions along the magnetic field the $\mathbf{v} \times \mathbf{B}$ force is zero, so the magnetic field has no effect on the $z-z$ element of the dielectric tensor. Following Stix [1992], the terms S and D can be decomposed into a sum and difference using the relations

$$S = \frac{1}{2}(R + L) \quad \text{and} \quad D = \frac{1}{2}(R - L), \quad (4.4.10)$$

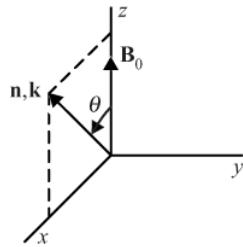


Fig. 4.14 The coordinate system used for analyzing wave propagation in a cold magnetized plasma.

where R and L are defined by

$$R = 1 - \sum_s \frac{\omega_{ps}^2}{\omega(\omega + \omega_{cs})} \quad \text{and} \quad L = 1 - \sum_s \frac{\omega_{ps}^2}{\omega(\omega - \omega_{cs})}. \quad (4.4.11)$$

As will be shown, the equation R is associated with a right-hand polarized mode, and the equation L is associated with a left-hand polarized mode. Note that the sign of the cyclotron frequency in equation (4.4.8) is determined by the sign of the charge (i.e., negative for electrons and positive for positively charged ions).

To obtain the dispersion relation, it is convenient to express the homogeneous equation (4.2.9) in matrix form. Without loss of generality we can rotate the coordinate system so that the index of refraction vector \mathbf{n} (which is parallel to \mathbf{k}) is in the x, z plane at an angle θ relative to the magnetic field \mathbf{B}_0 as shown in Figure 4.14, so that

$$\mathbf{n} = (n \sin \theta, 0, n \cos \theta). \quad (4.4.12)$$

Computing the various cross-products in equation (4.2.9) and organizing the results in matrix form, the homogeneous equation for the electric field can be written

$$\begin{bmatrix} -n^2 \cos^2 \theta & 0 & n^2 \sin \theta \cos \theta \\ 0 & -n^2 & 0 \\ n^2 \sin \theta \cos \theta & 0 & -n^2 \sin^2 \theta \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} + \begin{bmatrix} S & -iD & 0 \\ -iD & S & 0 \\ 0 & 0 & P \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0 \quad (4.4.13)$$

which reduces to

$$\begin{bmatrix} S - n^2 \cos^2 \theta & -iD & n^2 \sin \theta \cos \theta \\ iD & S - n^2 & 0 \\ n^2 \sin \theta \cos \theta & 0 & P - n^2 \sin^2 \theta \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (4.4.14)$$

Following the procedure in the previous section, the homogeneous equation has a non-trivial solution if and only if the determinant of the matrix is zero, which gives the dispersion relation

$$\begin{aligned} D(\mathbf{n}, \omega) &= n^2 \sin \theta \cos \theta [-(S - n^2)n^2 \sin \theta \cos \theta] \\ &\quad + [P - n^2 \sin^2 \theta][(S - n^2)(S - n^2 \cos^2 \theta) - D^2] = 0. \end{aligned} \quad (4.4.15)$$

Note that the n^6 terms cancel. This is fortunate because otherwise the dispersion relation would be a cubic in n^2 , which would be very difficult to solve. Using the identity $S^2 - D^2 = RL$ the dispersion relation can be simplified to the form

$$D(\mathbf{n}, \omega) = An^4 - Bn^2 + RLP = 0, \quad (4.4.16)$$

where $A = S \sin^2 \theta + P \cos^2 \theta$ and $B = RL \sin^2 \theta + PS(1 + \cos^2 \theta)$. This equation is a quadratic in n^2 and can be solved using the quadratic formula, which gives

$$n^2 = \frac{B \pm F}{2A}, \quad (4.4.17)$$

where the term F can be written in the positive definite form

$$F^2 = (RL - PS)^2 \sin^4 \theta + 4P^2 D^2 \cos^2 \theta. \quad (4.4.18)$$

This form is useful because it shows that F must always be real. Since A and B are real, it follows that the index of refraction must be either purely real ($n^2 > 0$), which corresponds to a propagating wave, or purely imaginary ($n^2 < 0$), which corresponds to an evanescent wave. A complex index of refraction, with non-zero real and imaginary parts, cannot occur because such an index of refraction would imply absorption of wave energy by the plasma. In the absence of collisions, energy absorption cannot occur in a cold plasma because there is no mechanism for converting the ordered motions associated with the wave into random thermal motions.

An equivalent form of the dispersion relation can be obtained by multiplying the term RLP in equation (4.4.16) by $\sin^2 \theta + \cos^2 \theta (= 1)$ and sorting out the $\sin^2 \theta$ and $\cos^2 \theta$ terms. The result is

$$\tan^2 \theta = \frac{-P(n^2 - R)(n^2 - L)}{(S n^2 - RL)(n^2 - P)}. \quad (4.4.19)$$

This relation is sometimes referred to as the “tangent” form of the dispersion relation, and is useful in some circumstances.

4.4.1 Propagation parallel to the magnetic field

For propagation parallel to the magnetic field, $\theta = 0$, equation (4.4.19) shows that the dispersion relation has three roots: $P = 0$, $n^2 = R$, and $n^2 = L$. To understand the nature of

the waves associated with each of these roots, it is necessary to examine the homogeneous equation for the electric field.

For $\theta = 0$ the homogeneous equation (4.4.14) becomes

$$\begin{bmatrix} S - n^2 & -iD & 0 \\ -iD & S - n^2 & 0 \\ 0 & 0 & P \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (4.4.20)$$

The electric field eigenvectors associated with each of the three roots are as follows:

$$P = 0, \quad \tilde{\mathbf{E}} = (0, 0, E_0), \quad (4.4.21)$$

$$n^2 = R, \quad \tilde{\mathbf{E}} = (E_0, iE_0, E_0), \quad (4.4.22)$$

$$n^2 = L, \quad \tilde{\mathbf{E}} = (E_0, -iE_0, E_0). \quad (4.4.23)$$

Since the electric field eigenvector for the first root, $P = 0$, is parallel to the z axis, and therefore to \mathbf{k} , this root corresponds to a longitudinal wave. This is the same longitudinal electrostatic mode described by equation (4.3.15) for a cold unmagnetized plasma. The magnetic field has no effect on this mode because the particles oscillate along the magnetic field and the $\mathbf{v} \times \mathbf{B}$ force vanishes. As can be seen from equation (4.4.9), the condition $P = 0$ implies that the oscillation occurs at the plasma frequency, $\omega = \pm\omega_p$.

Since the electric field eigenvectors for the second and third roots, $n^2 = R$ and $n^2 = L$, are perpendicular to the wave vector, \mathbf{k} , which for $\theta = 0$ is along the z axis, these roots correspond to transverse waves. It is easily verified that these modes are electromagnetic, since Faraday's law, $\tilde{\mathbf{B}} = \mathbf{k} \times \tilde{\mathbf{E}}/\omega$, shows that the wave magnetic field is non-zero. Since the field geometry is such that $\mathbf{k} \cdot \tilde{\mathbf{E}} = 0$, Gauss' law, $\tilde{\rho}_q = \epsilon_0(i\mathbf{k}) \cdot \tilde{\mathbf{E}}$, shows that there are no charge density fluctuations associated with these modes. When computing the eigenvectors, note that for $n^2 = R$, one has $S - n^2 = -(1/2)(R - L) = -D$, whereas for $n^2 = L$, one has $S - n^2 = D$. Thus, the only difference between the eigenvectors for these two modes is the sign of \tilde{E}_y . The $+i$ in the \tilde{E}_y component for the $n^2 = R$ mode indicates a $\pi/2$ phase shift with respect to the \tilde{E}_x component. It is easily verified that this phase shift corresponds to a right-hand sense of rotation with respect to the static magnetic field. For example, if E_0 is a real number then at $z = 0$

$$E_x = \operatorname{Re}\{E_0 e^{-i\omega t}\} = E_0 \cos \omega t, \quad (4.4.24)$$

and

$$E_y = \operatorname{Re}\{iE_0 e^{-i\omega t}\} = E_0 \sin \omega t, \quad (4.4.25)$$

which represents a wave rotating in the right-hand sense with respect to the static magnetic field. Similarly, the $-i$ in the E_y component of the electric field eigenvector for the $n^2 = L$ mode indicates a left-hand polarized wave. Therefore, the R and L notation for the two modes gives the sense of rotation (right-hand and left-hand) with respect to the externally

imposed static magnetic field. Note that this convention does *not* correspond to the convention used in optics. In optics the polarization is defined with respect to the wave vector, \mathbf{k} , rather than the magnetic field vector \mathbf{B}_0 .

It is useful next to consider the frequency dependence of the index of refraction for the R and L modes. For the R mode the index of refraction is given by

$$n^2 = R = 1 - \sum_s \frac{\omega_{ps}^2}{\omega(\omega + \omega_{cs})}. \quad (4.4.26)$$

A plot of n^2 as a function of frequency is shown in Figure 4.15 for a plasma consisting of electrons and two types of positively charged ions, with cyclotron frequencies ω_{ce} , ω_{c1} and ω_{c2} . Note that since ω_{ce} is negative for electrons, the function $n^2 = R$ goes to infinity at $\omega = |\omega_{ce}|$. A condition where the index of refraction goes to infinity is called a *resonance*. In this case the resonance is called the electron cyclotron resonance. Since electrons rotate around the static magnetic field in the same (right-hand) sense as the rotation of the wave electric field, for frequencies near the electron cyclotron frequency the wave interacts very strongly with the electrons. A comparable resonant interaction does not occur at the ion cyclotron frequencies ω_{c1} and ω_{c2} , since the wave field and the ions have opposite senses of rotation.

At a frequency somewhat above the electron cyclotron frequency the index of refraction goes to zero. A condition where the index of refraction goes to zero is called a *cutoff*. The frequency where this cutoff occurs is called the $R = 0$ cutoff. To a very good approximation, ion motions can be ignored at these high frequencies. Omitting the ion terms in equation (4.4.26), it is easy to show by setting $R = 0$ that the cutoff frequency is given by

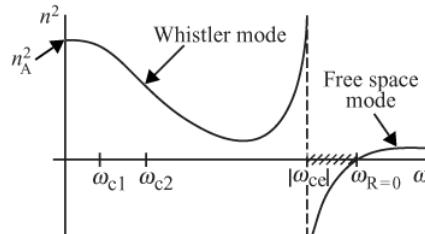


Fig. 4.15 | A plot of $n^2 = R$ as a function of frequency.

$$\omega_{R=0} = \frac{|\omega_{ce}|}{2} + \sqrt{\left(\frac{|\omega_{ce}|}{2}\right)^2 + \omega_{pe}^2}. \quad (4.4.27)$$

Since n^2 is negative between $|\omega_{ce}|$ and $\omega_{R=0}$, right-hand polarized waves cannot propagate in this range of frequencies, since the index of refraction is imaginary. The wave becomes *evanescent*. The evanescent region is indicated by the hatched lines in Figure 4.15. The branch of the dispersion relation below $|\omega_{ce}|$ is called the *whistler mode*, after a type of magnetospheric radio wave called a whistler, which will be discussed shortly. The branch of the dispersion relation above $\omega_{R=0}$ is called the free space mode because at high frequencies, this branch becomes the right-hand polarized free space electromagnetic mode.

Note that as the frequency approaches infinity the index of refraction approaches one, its free space value.

As the wave frequency approaches zero, the function R appears to diverge because of the $1/\omega$ terms in equation (4.4.26). However, careful analysis shows that the coefficients of the $1/\omega$ terms cancel if the plasma is electrically neutral ($\sum_i n_{0i} = n_{0e}$). In the limit of zero frequency, it can be shown that the index of refraction approaches a constant value called the Alfvén index of refraction, n_A , which is given by

$$n_A^2 = 1 + \sum_s \frac{\omega_{ps}^2}{\omega_{cs}^2}. \quad (4.4.28)$$

Waves obeying this dispersion relation are called Alfvén waves, after Alfvén [1942] who first predicted their existence. A more detailed discussion of Alfvén waves is given in Chapter ??.

The index of refraction for the left-hand polarized mode is given by

$$n^2 = L = 1 - \sum_s \frac{\omega_{ps}^2}{\omega(\omega - \omega_{cs})}. \quad (4.4.29)$$

The frequency dependence of the L mode is similar to the R mode, except that the resonances are now at the ion cyclotron frequencies rather than at the electron cyclotron frequency (again assuming positively charged ions). These resonances are called ion cyclotron resonances. They occur because positively charged ions rotate around the static magnetic field in the same left-hand sense as the rotation of the wave electric field, thereby causing strong interactions near the ion cyclotron frequencies. A plot of n^2 as a function of frequency is shown in Figure 4.16. The frequency variation consists of an alternating series of resonances and cutoffs that separate regions of propagation and evanescence. The evanescent regions are indicated by hatched lines. The branches below each of the ion cyclotron frequencies are called the ion cyclotron modes. As the frequency approaches zero, the index of refraction approaches the Alfvén index of refraction, just as in the case of the R mode. At high frequencies the index of refraction approaches the free space value ($n \rightarrow 1$ as $(\omega \rightarrow \infty)$). This branch is called the left-hand polarized free space mode. The low-frequency cutoff of the free space branch can be obtained by setting $L = 0$ in equation (4.4.11), and is called the $L = 0$ cutoff. Usually ion motions can be ignored when computing this cutoff, in which case the cutoff frequency is given by

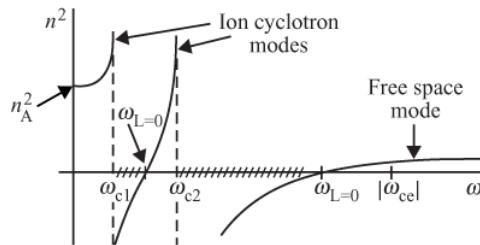


Fig. 4.16 A plot of $n^2 = L$ as a function of frequency.

$$\omega_{L=0} = \frac{|\omega_{ce}|}{2} + \sqrt{\left(\frac{\omega_{ce}}{2}\right)^2 + \omega_{pe}^2}. \quad (4.4.30)$$

This approximation fails at very low densities, since in this limit $\omega_{L=0}$ becomes comparable to the ion cyclotron frequencies. The cutoff frequency must then be calculated numerically. Cutoff frequencies also occur between the ion cyclotron frequencies. These cutoffs are called the ion-ion cutoffs, and usually must be computed numerically.

Numerous examples of the above modes of propagation exist in both laboratory and space plasmas. Some of these examples are discussed below.

4.4.1.1 Whistlers

In 1918 Barkhausen [1919], while listening to signals from an antenna connected to a rudimentary vacuum tube amplifier, reported hearing unusual whistling tones that decrease in frequency with increasing time. These signals came to be known as whistlers. Over thirty years later, in 1953, Storey [1953] explained the origin of whistlers. He showed that whistlers are produced by lightning and that the whistling tone is caused by the dispersive propagation of the lightning signal as it travels along a magnetic field line from one hemisphere to the other in the right-hand polarized mode now known as the whistler mode.

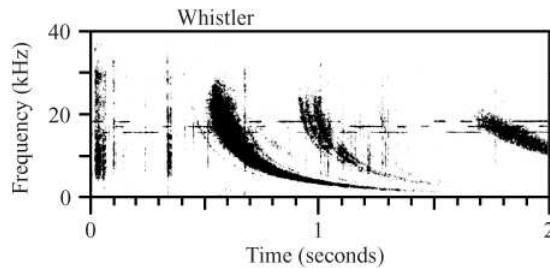


Fig. 4.17 A frequency–time spectrogram of several whistlers [adapted from Helliwell, 1965].

A frequency–time spectrogram of several whistlers detected on the ground is shown Figure 4.17. Dark regions in the spectrogram indicate higher intensities. According to currently accepted ideas, whistlers are guided almost exactly along the magnetic field line ($\theta = 0$) by small magnetic field aligned density irregularities that exist in the magnetized plasma surrounding the Earth called the magneto-sphere. The guiding of the wave is thought to be similar to the guiding of light in an optical fiber. Later, we will show that even in the absence of such field aligned irregularities, the whistler energy still follows the magnetic field to a good approximation because of the anisotropy introduced by the magnetic field. The whistling character of the signal arises because of the dispersive nature of the propagation. At the low (audio) frequencies where whistlers are typically observed, the higher frequencies propagate faster than lower frequencies. Since the lightning stroke simultaneously emits a broad range of frequencies, the higher frequencies arrive first, thereby dispersing the signal from the lightning stroke into a tone that decreases in frequency with increasing time. Because the wave can reflect at the base of the ionosphere, whistlers sometimes

bounce back and forth along the magnetic field line from one hemisphere to the other. A periodic sequence of whistlers is then observed, with each successive whistler having proportionally more dispersion.

For the conditions that exist in the Earth's magnetosphere, the index of refraction of the whistler mode is usually much greater than one, so the first term in equation (4.4.26) can be discarded. Since the frequencies of interest are usually well above the ion cyclotron frequencies, the ion terms can also be discarded. The index of refraction is then given by

$$n^2 = \frac{\omega_p^2}{\omega(\omega_c - \omega)}, \quad (4.4.31)$$

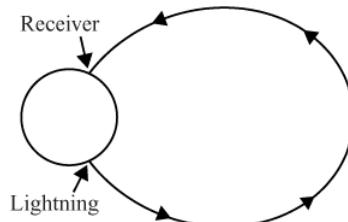


Fig. 4.18 Whistlers observed on the ground are guided along the magnetic field line from one hemisphere to the other by small field aligned density irregularities.

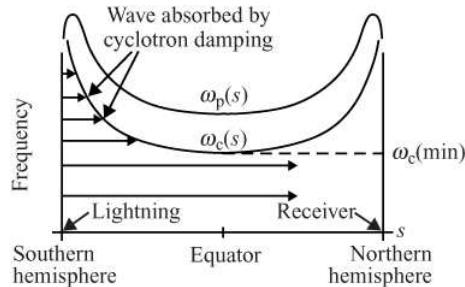


Fig. 4.19 A representative plot of the electron plasma frequency, $\omega_p(s)$, and the cyclotron frequency, $\omega_c(s)$, as a function of the path length, s , along the magnetic field line.

where, for notational convenience, we write ω_p for ω_{pe} and ω_c for $|\omega_{ce}|$. It is easy to show that the group velocity is given by

$$v_g = 2c \frac{\omega^{1/2}(\omega_c - \omega)^{3/2}}{\omega_c \omega_p}. \quad (4.4.32)$$

The travel time $t(\omega)$ for a given frequency ω is then given by

$$t(\omega) = \frac{1}{2c\omega^{1/2}} \int \frac{\omega_p \omega_c}{(\omega_c - \omega)^{3/2}} ds, \quad (4.4.33)$$

where the integration is carried out along the magnetic field line from the lightning source

to the receiver, as shown in Figure 4.18. In the above equation both ω_c and ω_p are functions of the path length s along the magnetic field line. Representative plots of $\omega_c(s)$ and $\omega_p(s)$ as a function of the path length, s , from one hemisphere to the other are shown in Figure 4.19. Note from equation (4.4.33) that the travel time goes to infinity at zero frequency and at the minimum electron cyclotron frequency along the path, $\omega_c(\min)$. Frequencies above $\omega_c(\min)$ are absorbed at the electron cyclotron frequency during the upgoing portion of the ray path by a process called cyclotron damping (see Chapter ??) and cannot reach the opposite hemisphere.

Since the travel time depends on the electron density (via the plasma frequency, ω_p) it is obvious that the whistler dispersion gives information on the electron density. Since ω_c is known, the integral equation for $t(\omega)$ can in principle be inverted to give $\omega_p(s)$. However, in practice small errors in $t(\omega)$ usually lead to large errors in the computed electron density, so certain simplifying assumptions are usually made. At frequencies well below the electron cyclotron frequency, $\omega \ll \omega_c$, where whistlers are usually observed, it is easy to show that equation (4.4.33) simplifies to $t(f) = D_W/f^{1/2}$, where $f = \omega/2\pi$ and

$$D_W = \frac{1}{2c} \int \frac{f_p}{f_c^{1/2}} ds, \quad (4.4.34)$$

is called the dispersion. By measuring the travel time as a function of frequency it is easy to determine the dispersion. From the above equation one can see that the dispersion is controlled by the line integral of $(n_e/B)^{1/2}$ along the magnetic field line. Since the magnetic field B reaches a minimum near the equator, the dispersion is mainly controlled by the electron density near the equator. In the pre-space era, before direct in situ measurements were possible, whistlers provided the only method of measuring the electron density at high altitudes, above the Earth's ionosphere.

4.4.1.2 Ion cyclotron whistlers

When plasma wave receivers were first flown on Earth orbiting spacecraft, a new type of whistler was discovered that cannot be observed from the ground. This type of whistler propagates in the left-hand polarized ion cyclotron mode, and is called an ion cyclotron whistler [Gurnett *et al.*, 1965]. Ion cyclotron whistlers occur immediately after an upward propagating right-hand polarized whistler and consist of one or more slowly rising tones lasting several seconds. The frequencies of the tones asymptotically approach the ion cyclotron frequencies of the local plasma. A frequency–time spectrogram showing a pair of ion cyclotron whistlers observed at an altitude of about 1300 km over a lightning storm is shown in Figure 4.20. The long tones occur because the group velocity of the ion cyclotron mode goes to zero at the ion cyclotron frequencies, H^+ and He^+ in this case. Ion cyclotron whistlers exhibit an unusual feature called polarization reversal. If plots of $n^2 = R$ and $n^2 = L$ versus frequency are superposed on the same diagram, as in Figure 4.21, the plots cross at a series of frequencies called the crossover frequencies, ω_x . The crossover frequencies are solutions of the equation $D = 0$, which can be written in the form

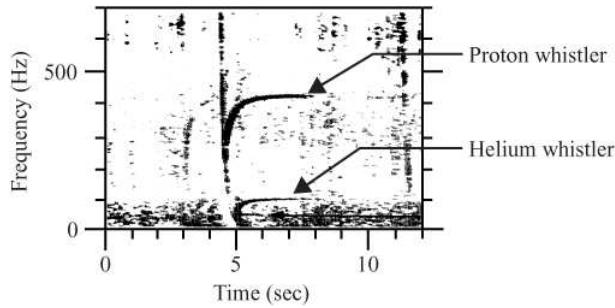


Fig. 4.20 A frequency–time spectrogram of two ion cyclotron whistlers detected in the ionosphere over a lightning storm.

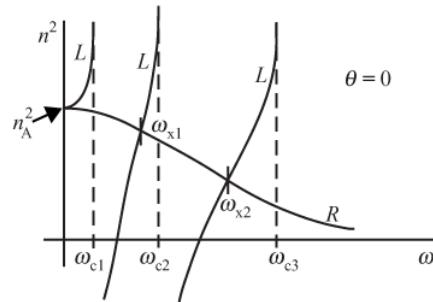


Fig. 4.21 Superimposed plots of $n^2 = R$ and $n^2 = L$. The plots intersect at the crossover frequencies, ω_{x1} and ω_{x2} .

$$D = \sum_s \frac{\omega_{ps}^2 \omega_{cs}}{\omega_x(\omega_x^2 - \omega_{cs}^2)} = 0. \quad (4.4.35)$$

Inspection of the above equation shows that one crossover frequency occurs between each pair of adjacent ion cyclotron frequencies.

Because the index of refraction curves for $\theta = 0$ intersect at the crossover frequencies, it is not immediately obvious how the two branches of the dispersion relation are connected at these intersections. This question can be answered by numerically computing the index of refraction slightly away from $\theta = 0$ using the general dispersion relation (4.4.17). The R and L branches are found to connect as shown in Figure 4.22. The polarization of a given branch is determined by the ratio of \tilde{E}_y to \tilde{E}_x , which from equation (4.4.20) is given by

$$\frac{\tilde{E}_y}{\tilde{E}_x} = \frac{-iD}{S - n^2}. \quad (4.4.36)$$

From the above equation one can see that the polarization, \tilde{E}_y/\tilde{E}_x , changes sign as D goes through zero (i.e., at the crossover frequencies). Thus, for a wave propagating in an inhomogeneous plasma, the polarization reverses whenever the crossover frequency moves across the wave frequency.

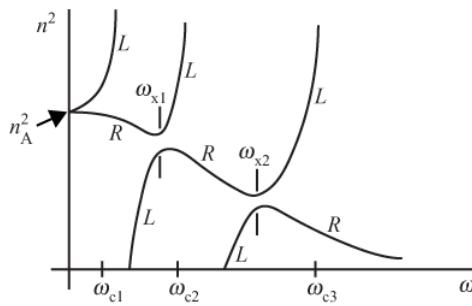


Fig. 4.22 | The index of refraction of the R and L modes at a small but non-zero wave normal angle.

To illustrate in more detail how this polarization reversal occurs, it is useful to consider how the cyclotron frequencies and crossover frequencies vary as a function of altitude. Three primary ion species, O^+ , He^+ and H^+ are present in the Earth's ionosphere. The corresponding ion cyclotron frequencies are ω_{c1} , ω_{c2} , and ω_{c3} . Because of their different molecular weights these ions are separated into distinct height bands, with O^+ dominant at low altitudes, He^+ at intermediate altitudes, and H^+ at high altitudes. These concentration variations cause the crossover frequency to vary with altitude as shown in Figure 4.23. One can see from this plot that an upward propagating right-hand polarized wave, R , from a lightning stroke at frequency ω_2 is converted to a left-hand polarized ion cyclotron wave, L , at crossover frequency ω_{x2} . This wave becomes a proton (H^+) whistler. At frequency ω_1 a similar polarization reversal also occurs at crossover frequency ω_{x1} . This wave becomes a helium (He^+) whistler.

After the upgoing right-hand polarized wave from the lightning stroke has been converted to the left-hand polarized ion cyclotron mode, it continues to propagate upward, approaching closer and closer to the ion cyclotron frequency. The wave energy is eventually absorbed at the ion cyclotron frequency by a process called cyclotron damping (see Chapter ??). The above geometry, where a wave approaches resonance in a region of decreasing magnetic field, is called a magnetic beach, in analogy with the dissipation of ocean waves propagating into a beach. The magnetic beach geometry has been used for heating ions in fusion machines and is called ion cyclotron heating.

4.4.1.3 Faraday rotation

As can be seen by comparing Figures 4.15 and 4.16, both the left-hand and right-hand polarized free space modes can propagate at frequencies above the right-hand cutoff, $\omega_{R=0}$. In this frequency range, above the right-hand cutoff, a wave of any arbitrary polarization (linear, elliptical, or circular) can be constructed from a linear superposition of the right-hand and left-hand polarized waves. However, unlike in free space, in a plasma the phase velocities of the right-hand and left-hand polarized modes are not the same. The difference in the phase velocities causes the electric field to rotate as the wave propagates through the

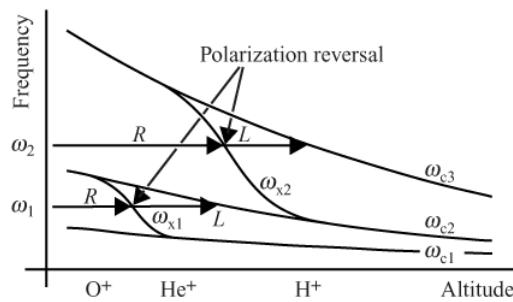


Fig. 4.23 A representative plot of the crossover frequencies ω_{x1} and ω_{x2} as a function of altitude.

plasma. This effect is called Faraday rotation, and occurs in a variety of astrophysical and laboratory situations.

To illustrate the basic effect, consider the special case of a wave that is constrained to be linearly polarized along the x axis at $z = 0$, with \mathbf{k} and \mathbf{B}_0 aligned along the z axis. This boundary condition can be satisfied by superposing left-hand and right-hand circularly polarized waves of the form

$$E_y = \frac{1}{2} E_0 [\sin(k_L z - \omega t) - \sin(k_R z - \omega t)] \quad (4.4.37)$$

and

$$E_x = \frac{1}{2} E_0 [\cos(k_L z - \omega t) - \cos(k_R z - \omega t)], \quad (4.4.38)$$

where K_R and k_L are the wave numbers of the right-hand and left-hand polarized modes. It is easily verified that the electric field is aligned along the x axis at $z = 0$. Using well-known trigonometric identities the above two equations can be rewritten in the form

$$E_y = E_0 \sin \left[\frac{1}{2} (k_L - k_R) z \right] \cos \left[\frac{1}{2} (k_L + k_R) z - \omega t \right] \quad (4.4.39)$$

and

$$E_x = E_0 \cos \left[\frac{1}{2} (k_L - k_R) z \right] \cos \left[\frac{1}{2} (k_L + k_R) z - \omega t \right]. \quad (4.4.40)$$

Taking the ratio of E_y to E_x , it is easy to show that

$$\frac{E_y}{E_x} = \tan \psi, \text{ where } \psi = \frac{1}{2} (k_L - k_R) z. \quad (4.4.41)$$

The above equation shows that the wave remains linearly polarized at all points along the z axis, and that the plane of polarization rotates through an angle ψ relative to the plane of polarization at $z = 0$. Since ψ increases linearly with z , the electric field vector lies on a

helix as shown in Figure 4.24. A similar rotation also occurs for an elliptically polarized wave.

If the wave frequency is much greater than both the electron cyclotron frequency and the electron plasma frequency, a simple formula can be derived for the Faraday rotation angle. Dropping the ion terms in equations (4.4.26) and (4.4.29) and assuming that $\omega \gg \omega_c$ and $\omega \gg \omega_p$, the indices of refraction of the R and L modes can be approximated by

$$n_R \approx 1 - \frac{1}{2} \frac{\omega_p^2}{\omega^2} \left(1 + \frac{\omega_c}{\omega} \right) \quad (4.4.42)$$

and

$$n_L \approx 1 - \frac{1}{2} \frac{\omega_p^2}{\omega^2} \left(1 - \frac{\omega_c}{\omega} \right). \quad (4.4.43)$$

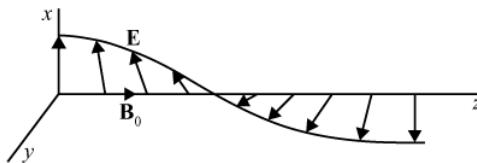


Fig. 4.24 The electric field of a linearly polarized free space electromagnetic wave rotates as the wave propagates through a magnetized plasma.

Converting n to k using the relation $k = \omega n/c$, it is easy to show using equation (4.4.41) that the Faraday rotation angle is given by

$$\psi \approx \frac{1}{2c} \left(\frac{\omega_p^2 \omega_c}{\omega^2} \right) z. \quad (4.4.44)$$

This equation shows that ψ increases in direct proportion to the product of the electron density and the magnetic field strength. If the electron density and magnetic field strength vary along the ray path, the rotation angle is given by the integral

$$\psi = \frac{1}{2c} \left(\frac{e^3}{\epsilon_0 m_e^2} \right) \frac{1}{\omega^2} \int n_e B_z d\ell, \quad (4.4.45)$$

where the integration is performed along the ray path. If the magnetic field is at an angle to the direction of propagation, it can be shown that equation (4.4.45) is still valid, provided B_z is taken to be the component of the magnetic field along the direction of propagation.

Measurements of the Faraday rotation angle can be used to provide information on the electron density or magnetic field strength along the ray path. Since the rotation angle involves the product of the electron density and the magnetic field, one of these parameters must be known to get useful results. For example, in a laboratory plasma, the magnetic field is usually known, so the line integral of the electron density can be obtained by measuring the Faraday rotation angle. In practice, this is done by generating a linearly polarized wave

with a microwave transmitter on one side of a plasma chamber and measuring the polarization angle with a suitable antenna on the other side of the chamber. Similarly, ground-based measurements of the Faraday rotation angle of signals from a satellite transmitter can be used to give information on the electron density in the Earth's ionosphere.

Because of the $1/\omega^2$ dependence in expression (4.4.45), it is not necessary to know the polarization of the source if measurements are made at two or more frequencies. Such multi-frequency Faraday rotation measurements have been used to provide estimates of the magnetic field strength in the interstellar medium. By measuring the Faraday rotation of signals from a distant radio source, such as a pulsar, the integral $\int n_e B_z d\ell$ can be obtained. The problem is then to unscramble the n_e and B_z dependences. For a pulsar, the line integral $\int n_e d\ell$ along the ray path from the pulsar to the Earth can be determined from the dispersion of the pulsar signal. The ratio of the two integrals can then be used to give a measure of the average magnetic field component along the propagation path. The sign of B_z can also be determined from the sense of the Faraday rotation. This technique has been used by Manchester and Taylor [1977] to determine the strength and orientation of the magnetic field in our galaxy.

4.4.2 Propagation perpendicular to the magnetic field

For propagation perpendicular to the magnetic field, $\theta = \pi/2$, equation (4.4.19) shows that the dispersion relation has two roots: $n^2 = P$ and $n^2 = RL/S$. The electric field eigenvectors associated with these two roots can be obtained from the homogeneous equation (4.4.14), which for $\theta = \pi/2$ becomes

$$\begin{bmatrix} S & -iD & 0 \\ iD & S - n^2 & 0 \\ 0 & 0 & P - n^2 \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (4.4.46)$$

The electric field eigenvectors associated with the two roots are as follows:

$$n^2 = P, \quad \tilde{\mathbf{E}} = (0, 0, E_0); \quad (4.4.47)$$

and

$$n^2 = \frac{RL}{S}, \quad \tilde{\mathbf{E}} = \left(\frac{iD}{S} E_0, E_0, 0 \right). \quad (4.4.48)$$

The root $n^2 = P$ is identical to the dispersion relation for transverse waves in a cold unmagnetized plasma, as discussed in Section 4.3.1. The static magnetic field has no effect on this mode because the particle motion is parallel to the magnetic field, and hence there is no $\mathbf{v} \times \mathbf{B}_0$ force. The relative orientations of the \mathbf{k} vector, the wave electric and magnetic fields, and the static magnetic field are shown in Figure 4.25. Since there is no dependence on the magnetic field strength, this mode of propagation is called the ordinary (O) mode. Note that the electric field in this case is linearly polarized.

The root $n^2 = RL/S$ is more complicated. For this mode the electric field consists of both

longitudinal (parallel to \mathbf{k}) and transverse (perpendicular to \mathbf{k}) components. Therefore, this mode has both electrostatic and electromagnetic characteristics. The relative orientations of the \mathbf{k} vector, the wave electric and magnetic fields, and the static magnetic field are shown in Figure 4.26. Since this mode of propagation depends on the magnetic field strength, this mode is called the extraordinary (X) mode.

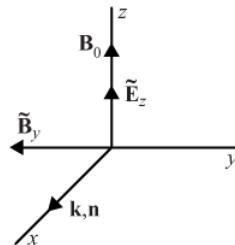


Fig. 4.25 The electric and magnetic field eigenvectors for the ordinary (O) mode.

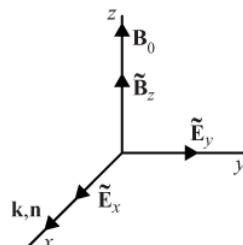


Fig. 4.26 The electric and magnetic field eigenvectors for the extraordinary (X) mode.

The frequency dependence of the index of refraction for the extraordinary mode is determined by the functions R , L and S . The functions R and L have already been described in Section 4.4.1. The function S , defined by equation (4.4.8), consists of alternating poles and zeros, with a pole at each cyclotron frequency. The index of refraction, $n^2 = RL/S$, has zeros at the zeros of R and L and resonances at the zeros of S . Plots of S and $n^2 = RL/S$ for a plasma consisting of electrons and two ion species are shown in Figure 4.27. The resonances at $S = 0$ are called hybrid resonances because they involve a combination (or hybrid) of electric and magnetic forces. From the electric field eigenvector, $\tilde{\mathbf{E}} = [(iD/S)E_0, E_0, 0]$, one can see that as S approaches zero the longitudinal component $\tilde{E}_x = (iD/S)E_0$ becomes arbitrarily large compared to the transverse component $\tilde{E}_y = E_0$. Therefore, near the hybrid resonances the wave becomes almost purely electrostatic, similar in many respects to electron plasma oscillations, except magnetic forces are now important.

As can be seen in Figure 4.27, a hybrid resonance is associated with each species present in the plasma. For a plasma consisting of electrons and two ion species, there are three hybrid resonances: one above the electron cyclotron frequency, one between the electron cyclotron frequency and the largest ion cyclotron frequency, and one between the two ion cyclotron frequencies. These resonances are called the upper hybrid resonance, ω_{UH} , the

lower hybrid resonance, ω_{LH} , and the ion hybrid resonance, ω_{IH} , respectively. Additional ion hybrid resonances occur if more ion species are present.

Since the upper hybrid resonance frequency is much greater than any of the ion cyclotron frequencies, the upper hybrid resonance mainly involves electron motions. The corresponding solution to $S = 0$ can be computed to a good approximation by omitting the ion terms in equation (4.4.8), which gives

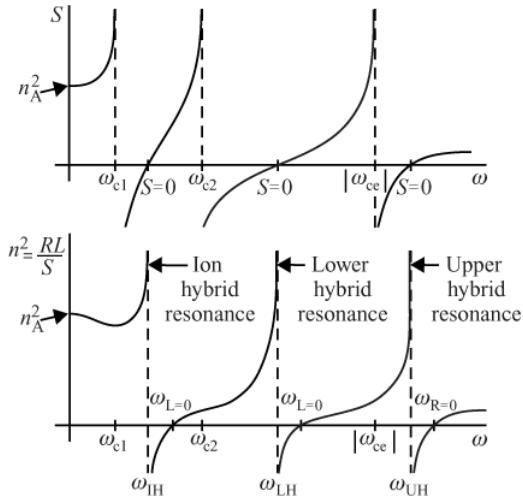


Fig. 4.27 Plots of S and $n^2 = RL/S$ as a function of frequency.

$$\omega_{UH} = \sqrt{\omega_p^2 + \omega_c^2}. \quad (4.4.49)$$

If the magnetic field is reduced to zero ($\omega_c \ll \omega_p$) the oscillation at the upper hybrid resonance reduces to the electron plasma oscillation found in an unmagnetized plasma.

Since the lower hybrid resonance frequency is between the electron cyclotron frequency and the largest ion cyclotron frequency, both the electron and ion terms must be included in equation (4.4.8) when solving for the corresponding root of $S = 0$. For simplicity, we consider a plasma consisting of electrons and only one ion species. In order to obtain a simple result one must make the further assumption that $\omega_{ci} \ll \omega_{LH} \ll |\omega_{ce}|$. With this assumption it can be shown that the lower hybrid resonance frequency is given to a good approximation by

$$\frac{1}{\omega_{LH}^2} = \frac{1}{\omega_{pi}^2} + \frac{1}{|\omega_{ce}|\omega_{ci}}. \quad (4.4.50)$$

Two further approximations, called the high density and low density limits, are often employed when using the above equation. In the high density limit, defined by $|\omega_{ce}|\omega_{ci} \ll \omega_{pi}^2$ (or equivalently $\omega_{ce}^2 \ll \omega_{pe}^2$), the first term on the right can be ignored, in which case the lower hybrid resonance frequency is given by the geometric mean of the electron and ion cyclotron frequencies

$$\omega_{LH} = \sqrt{|\omega_{ce}| \omega_{ci}}. \quad (4.4.51)$$

The assumption that $\omega_{LH} \ll |\omega_{ce}|$ is clearly satisfied by this solution. Both electron and ion motions are involved in the oscillation. In the low density limit, which is defined by $|\omega_{ce}| \omega_{ci} \gg \omega_{pi}^2$, the lower hybrid resonance frequency is given by

$$\omega_{LH} = \omega_{pi}. \quad (4.4.52)$$

In this case the electrons are held essentially motionless by the strong magnetic field. The ions then play the dominant role in the oscillation. These ion oscillations are exactly analogous to the electron plasma oscillations described in Section ??, except the ions move instead of the electrons.

The ion hybrid resonances involve motions of two or more ion species. An ion hybrid resonance occurs between each adjacent pair of ion cyclotron frequencies. It is not possible to obtain a simple analytic expression for the ion hybrid resonance frequencies because the corresponding roots of $S = 0$ involve equations of third (or higher) degree in ω^2 . These roots must be computed numerically.

Numerous examples of hybrid resonances exist in both laboratory and space plasmas. Some of these examples are discussed below.

4.4.2.1 Ionospheric sounders

Early satellite studies of the Earth's ionosphere utilized a technique called ionospheric sounding. Ionospheric sounding consists of transmitting a short pulse of radio waves at a frequency ω , and then recording the intensity of the signals reflected by the plasma [Franklin and Maclean, 1969]. After a suitable listening interval the transmitter frequency is increased by a small increment and the transmit/receive cycle is repeated. This cycle is repeated again and again until a complete scan is obtained of all of the characteristic frequencies of the plasma. Typically two types of reflected signals are detected: (1) echoes associated with remote reflections of electromagnetic waves by the ionosphere; and (2) electrostatic oscillations excited in the plasma near the spacecraft. The reflected signal intensity is typically plotted as a function of the transmitter frequency and the time delay after transmission of the pulse. This type of display is called an ionogram. An example of an ionogram obtained from an ionospheric sounder is shown in Figure 4.28. The darker regions indicate stronger signals. Usually the time delay is plotted vertically downward, since reflections are expected from the ionosphere, which is below the spacecraft. The lines labeled "free space (R, X)" and "free space (L, O)" are ionospheric echos of the right-hand and left-hand polarized free space modes. The intense vertical "spikes" in the ionogram are local electrostatic resonances. The vertical character of the spikes indicates that a long-duration oscillation is being excited. Very strong spikes occur at the electron plasma frequency, f_p , and the upper hybrid resonance frequency, F_{UH} . These correspond to the electrostatic oscillation at the electron plasma frequency and the upper hybrid resonance frequency. Strong responses are observed at these resonances because the group

velocity is zero at these frequencies, which causes the wave to remain in the vicinity of the spacecraft for a long time after the pulse is transmitted. Several other spike-like features are also evident in Figure 4.28, for example, at harmonics of the electron cyclotron frequency. These resonances involve electrostatic oscillations that require a finite temperature, and will be discussed in Chapter ??.

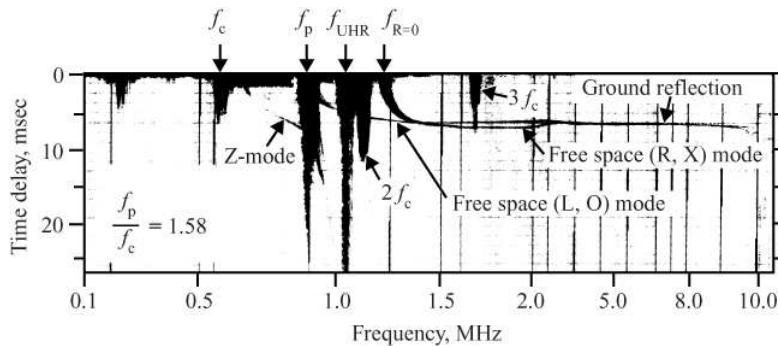


Fig. 4.28 An ionogram obtained from an ionospheric sounder on the Alouette 2 spacecraft at an altitude of 1542 km [adapted from Benson, 1982].

4.4.2.2 Thermal excitation of hybrid resonances

Early rocket and satellite measurements of waves in the Earth's ionosphere revealed narrowband emissions at both the upper and lower hybrid resonance frequencies [Walsh *et al.*, 1964; Brice and Smith, 1965]. These emissions are observed in the ionosphere and magnetosphere at geocentric radial distances of up to several Earth radii (R_E). An example of an emission at the upper hybrid resonance frequency is shown in Figure 4.29. The narrowband emission labeled “thermal emission at $f = f_{UH}$ ” is at the upper hybrid frequency, $F_{UH} = \sqrt{f_c^2 + f_p^2}$. In this case the electron cyclotron frequency is much less than the electron plasma frequency, $f_c \ll f_p$, so the upper hybrid frequency is essentially at the electron plasma frequency, which is proportional to the square root of the electron density, see equation (??). Therefore, the frequency of the upper hybrid emission line provides a direct measurement of the electron density. An electron density scale is given on the right-hand side of the spectrogram. The sharp peak in the electron density around the time of closest approach, indicated by the arrow marked “closest approach,” is due to the passage of the spacecraft through the ionosphere. These weak, nearly continuous emissions are believed to be caused by thermal excitation of electrostatic oscillations at the upper hybrid frequency. Occasionally, much more intense emissions are observed at the upper hybrid resonance frequency. These much stronger emissions are believed to be due to plasma instabilities, and will be discussed in Chapter ??.

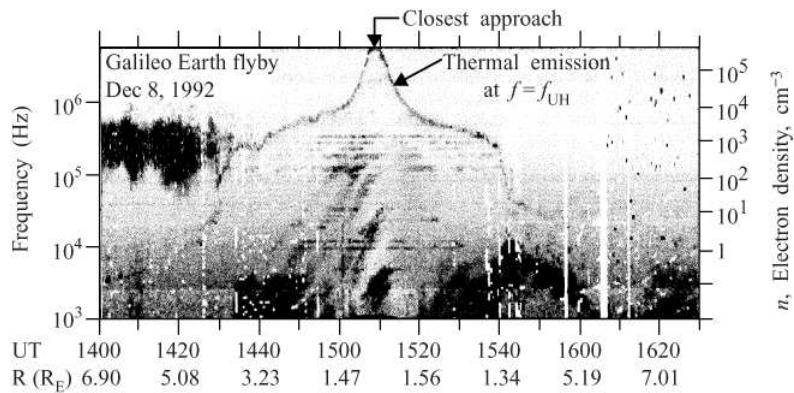


Fig. 4.29 A frequency–time spectrogram showing an emission line at the upper hybrid frequency during a close flyby of the Earth. UT is Universal Time.

4.4.2.3 Plasma heating via hybrid resonances

One of the basic challenges of controlled fusion is to heat the ions to very high temperatures, on the order of 10^8 K. Because of the greater mobility of electrons, simple Ohmic heating, such as by driving a current through the plasma, causes most of the energy to be absorbed by the electrons. What is needed is a mechanism that preferentially heats the ions. Because the lower hybrid and ion hybrid resonances involve primarily ion motions, it has been suggested that these resonances be used for ion heating. The basic idea is to transmit waves into the plasma perpendicular to the magnetic field at a frequency slightly below either the lower hybrid resonance frequency or the ion hybrid resonance frequency. If the density and magnetic field profiles are favorable, the wave can be brought into resonance at the appropriate hybrid resonance frequency inside the plasma, very similar to the magnetic beach geometry discussed for ion cyclotron whistlers. The wave energy is then absorbed at the resonance, heating the ions. The lower hybrid resonance is particularly favored for this purpose because in the high density limit, the resonance frequency $\omega_{LH} = \sqrt{\omega_{ci}|\omega_{ce}|}$ is easily controlled by the magnetic field. This method has been very successful in heating ions in laboratory fusion devices. Attempts have also been made to heat plasmas at the ion hybrid resonance frequency, with more limited success. In both cases the achieved heating rates are limited by the efficiency of the transmitter and the coupling of the radiated power into the plasma.

4.4.3 Oblique wave propagation

We next consider the propagation of waves at intermediate wave normal angles, neither parallel nor perpendicular to the magnetic field. The index of refraction at an arbitrary angle $n(\theta)$ is given by equation (4.4.17). Simple inspection of this equation shows that the index of refraction surface $n(\theta)$ has several obvious symmetries. Since the plasma has rotational symmetry around the magnetic field the index of refraction surface has a rotational symmetry around the \mathbf{B}_0 axis. The index of refraction surface also has a mirror symmetry

with respect to $\theta = \pi/2$. This symmetry arises because the index of refraction only involves terms of the form $\cos^2 \theta$ and $\sin^2 \theta$.

The topological shape of the index of refraction surface can be easily deduced from the general properties of equation (4.4.17). First, if the index of refraction goes to zero (i.e., a cutoff), then it must do so simultaneously for all angles. This result can be seen by examining equation (4.4.16). If $n = 0$ then either $R = 0, L = 0$, or $P = 0$. Since the condition for a cutoff does not involve the wave normal angle, then a cutoff must occur simultaneously for all wave normal angles. Second, if the index of refraction becomes infinite, then in the range $0 < \theta < \pi/2$ it can do so at only one angle, θ_{Res} , called the resonance cone angle. This result can be seen by examining equation (4.4.17). The index of refraction can go to infinity if and only if $A = 0$. However, the condition $A = 0$ can be satisfied if and only if S and P have opposite signs. If S and P have opposite signs then the index of refraction goes to infinity at a wave normal angle given by

$$\tan^2 \theta_{\text{Res}} = -\frac{P}{S}. \quad (4.4.53)$$

There are only two solutions of the above equation, one at θ_{Res} , and the other at $\pi - \theta_{\text{Res}}$. Therefore, if the index of refraction becomes infinite, then it must do so along two oppositely directed cones, the axes of which are aligned parallel to the magnetic field, as illustrated in Figure 4.30. Based on the above results, consider how the index of refraction must vary between $\theta = 0$ and $\theta = \pi/2$. Since the index of refraction cannot go to zero (unless it does so for all angles), and since n^2 can be either positive or negative at $\theta = 0$ and either positive or negative at $\theta = \pi/2$, it follows that there are only four possible ways for the index of refraction to vary between $\theta = 0$ and $\theta = \pi/2$. These four possibilities are illustrated in Figure 4.31. Of the four, only three give rise to real index of refraction surfaces. The fourth corresponds to an imaginary index of refraction at all angles, which is physically uninteresting. As shown in Figure 4.32, the index of refraction surfaces have the topological shapes of an ellipsoid, a hyperboloid of two sheets, and a hyperboloid of one sheet. The asymptotes of the hyperboloid-like surfaces are at the resonance cone angle.

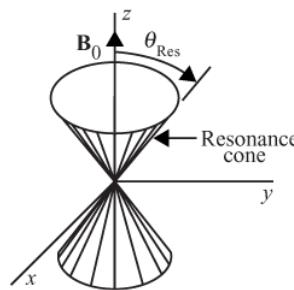


Fig. 4.30 The resonance cone.

The possible combinations of index of refraction surfaces that can occur when both modes of propagation are considered involve two basic conditions. First, except for degenerate cases at $\theta = 0$ and $\theta = \pi/2$, the two modes can never intersect. This follows from

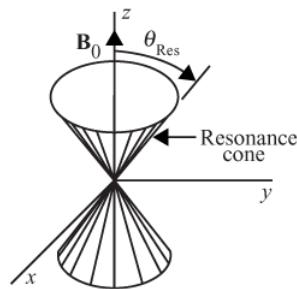


Fig. 4.31 The four possible plots of n^2 versus the wave normal angle, θ .

the fact that F in equation (4.4.18) is positive definite. The only exception occurs when the quantities $RL - PS$ and PD are simultaneously zero, which is physically unlikely. Second, only one of the two modes can have a resonance cone. This follows because if $n = \infty$ for one of the modes, then $A = 0$. Equation (4.4.16) then shows that the index of refraction of the other mode, $n^2 = RLP/B$, is finite. Considering the two possible signs of n^2 for each of the two modes at $\theta = 0$ and at $\theta = \pi/2$, a total of sixteen combinations occur. Of these, one is not of interest because both surfaces are imaginary, three represent redundant combinations, two violate the condition that the index of refraction surfaces cannot intersect, and four violate the condition that only one mode can have a resonance cone. After eliminating these possibilities, only six combinations survive. These six combinations are shown in Figure 4.33. The specific combination that occurs for any given set of plasma parameters is completely determined by the indices of refraction at $\theta = 0$ and $\theta = \pi/2$. Thus, using the n^2 plots in Figure 4.31, the topological shape of the index of refraction surfaces can be determined at any frequency, and for any combination of plasma parameters.

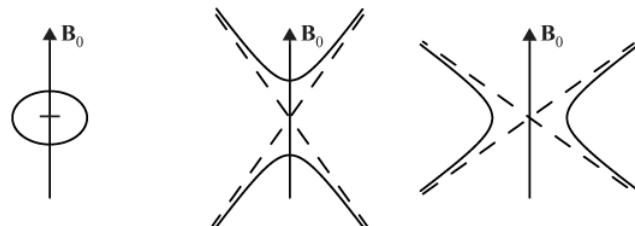


Fig. 4.32 The three possible types of index of refraction surfaces in a cold magnetized plasma.

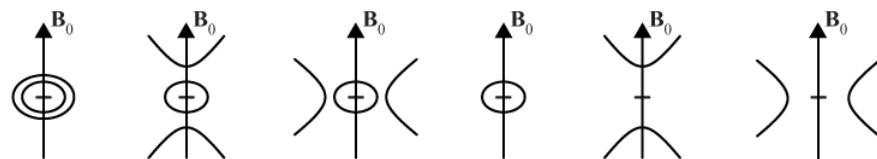


Fig. 4.33 The six possible combinations of index of refraction surfaces that can occur for the two electromagnetic modes of propagation in a cold magnetized plasma.

The above analysis shows that the index of refraction in a plasma is qualitatively different than in any ordinary dielectric media. Most striking is the highly anisotropic behavior of the index of refraction, especially the occurrence of resonance cones where the index of refraction goes to infinity. As an example that illustrates the highly unusual effects that can occur, we next describe the index of refraction of the whistler mode.

In order to obtain a mathematically tractable equation for the index of refraction of the whistler mode, it is necessary to make some simplifying assumptions. First, we assume that the ion terms can be ignored. Except for waves propagating very nearly perpendicular to the magnetic field, this assumption is normally good down to frequencies on the order of the ion cyclotron frequency. Second, we assume that the wave frequency and the electron cyclotron frequency are both much less than the electron plasma frequency, i.e., $\omega^2 \ll \omega_p^2$ and $\omega_c^2 \ll \omega_p^2$. This assumption is often called the high density approximation. With these limitations, the expressions for R , L , D , S , and P can be approximated as follows:

$$R \simeq \frac{-\omega_p^2}{\omega(\omega - \omega_c)}, L \simeq \frac{-\omega_p^2}{\omega(\omega + \omega_c)}, S \simeq \frac{-\omega_p^2}{\omega^2 - \omega_c^2}, D \simeq \frac{-\omega_p^2 \omega_c}{\omega(\omega^2 - \omega_c^2)}, P \simeq -\frac{\omega_p^2}{\omega^2}. \quad (4.4.54)$$

Noting that $RL = PS$, the quantities A , B and F in equation (4.4.17) simplify to the following

$$A \simeq \frac{-\omega_p^2(\omega^2 - \omega_c^2 \cos^2 \theta)}{\omega^2(\omega^2 - \omega_c^2)}, B \simeq \frac{2\omega_p^4}{\omega^2(\omega^2 - \omega_c^2)}, F \simeq 2 \frac{\omega_p^4 \omega_c}{\omega^3(\omega^2 - \omega_c^2)} \cos \theta. \quad (4.4.55)$$

Substituting expressions (4.4.54) and (4.4.55) into equation (4.4.17) and using the plus sign, which yields the whistler mode, the index of refraction is given by

$$n^2 = \frac{\omega_p^2}{\omega(\omega_c \cos \theta - \omega)}. \quad (4.4.56)$$

It is easy to check that this equation reduces to equation (4.4.31) when $\theta = 0$. A polar plot of $n(\theta)$ is shown in Figure 4.34. The index of refraction goes to infinity at the resonance cone angle, which in this case is given by $\cos \theta_{Res} = \omega/\omega_c$. Note the frequency dependence of the resonance cone angle. At high frequencies the resonance cone angle goes to zero at the electron cyclotron frequency. Therefore, the cone of wave normal angles within which propagation is allowed shrinks to zero as the wave frequency approaches the electron cyclotron frequency. Above the electron cyclotron frequency, the index of refraction is imaginary. At low frequencies, which is where whistlers are usually observed, the resonance cone angle approaches $\pi/2$. The index of refraction surface then has the shape shown in Figure 4.35. Since the group velocity is perpendicular to the index of refraction surface, it is evident from Figure 4.35 that the group velocity, v_g , is constrained to be within a small cone of angles relative to the magnetic field for all wave normal angles. It can be shown that the maximum angle ψ between the group velocity and the magnetic field is $\psi_{max} = \tan^{-1}(1/\sqrt{8}) = 19^\circ 28'$. This result was first shown by Storey [1953]. Therefore, even if there are insufficient density irregularities to maintain the wave normal angle at $\theta = 0^\circ$, whistlers are still guided along the magnetic field to a good approximation.

Furthermore, since the index of refraction goes to infinity at $\theta = \pi/2$, the wave normal angle cannot rotate through $\theta = \pi/2$. Thus, once a whistler starts up the magnetic field in one hemisphere, it cannot reverse its direction of motion until it reaches the base of the ionosphere in the opposite hemisphere. When ion effects are included it can be shown that the wave normal angle can rotate through $\theta = \pi/2$ at frequencies below the lower hybrid resonance frequency. This occurs because the index of refraction at $\theta = \pi/2$ is real for $\omega < \omega_{LHR}$.

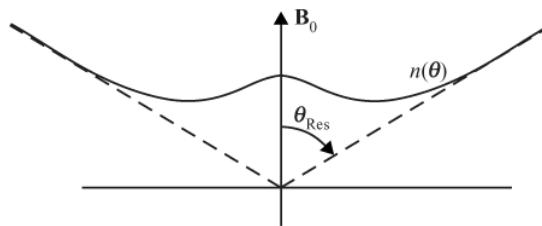


Fig. 4.34 The index of refraction for the whistler mode in the high density limit.

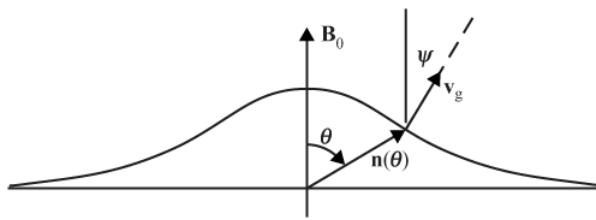


Fig. 4.35 The index of refraction for the whistler mode in the low frequency, high density limit.

4.4.4 The Clemmow-Mullaly-Allis (CMA) diagram

The dispersion relation derived in this chapter is based on the assumption that the plasma is homogeneous. The results obtained can be applied to inhomogeneous plasmas if the index of refraction varies sufficiently slowly. As a wave propagates through an inhomogeneous medium the plasma parameters can be expected to change. Because of these changes, a wave originating in a particular mode at one point in the plasma may or may not be able to reach some other region of the plasma. This issue is referred to as *accessibility*. Two points A and B are said to be accessible via a particular mode if a ray path can connect the two points. A necessary condition for accessibility is that the index of refraction be real and continuous at all points along some ray path between the two points. Continuity of the index of refraction is a necessary, but not a sufficient, condition for accessibility. It is not a sufficient condition because refraction may prevent a ray path from connecting the two points, even if the index of refraction is real and continuous between the two points. The effects of refraction are discussed in the next section.

To answer the question of whether the index of refraction is real and continuous between two points in a plasma, it is useful to consider a multi-dimensional parameter space called

a CMA diagram, after Clemmow and Mullaly [1955] and Allis [1959]. A CMA diagram consists of a diagram with one coordinate for each parameter of the plasma, such as n_s and B_0 . Within this parameter space a set of *bounding surfaces* is constructed defined by the equations

$$R = 0, \quad L = 0, \quad P = 0, \quad S = 0, \quad R = \infty \quad \text{and} \quad L = \infty.$$

These surfaces define *bounded volumes* in parameter space. Since R, L, S , and P can only change sign at a bounding surface, the topological form of the index of refraction surface remains the same throughout a bounded volume. Therefore, if the topological form is determined at one point in a bounded volume, it is known at all points in that bounded volume.

At a bounding surface the topological form of the index of refraction surface may or may not change. If the index of refraction surface changes from real to imaginary (i.e., non-propagating) at the bounding surface, then the transition across the bounding surface is called a *destructive* transition, since the wave cannot propagate into the region with the imaginary index of refraction. On the other hand, if the refractive index remains real and changes continuously upon crossing a bounding surface, then the transition is called a *non-destructive* transition, since a wave can in principle cross the bounding surface. All the bounded volumes connected by non-destructive transitions are possible accessible regions, subject only to the restrictions imposed by refraction.

4.4.4.1 The CMA diagram for a plasma of electrons and immobile ions

To illustrate these concepts we construct the CMA diagram for a plasma consisting of electrons and a fixed background of immobile ions. The assumption of immobile ions basically implies that the frequencies must be sufficiently high that ion motions are unimportant, i.e., well above the ion cyclotron frequencies. Dropping the ion terms, the expressions for R, L, S , and P become

$$R = 1 - \frac{\omega_p^2}{\omega(\omega - \omega_c)}, \quad L = 1 - \frac{\omega_p^2}{\omega(\omega + \omega_c)}, \quad S = 1 - \frac{\omega_p^2}{\omega^2 - \omega_c^2}, \quad P = 1 - \frac{\omega_p^2}{\omega^2}. \quad (4.4.57)$$

To proceed further, it is convenient to define the normalized, dimensionless parameters $Y = \omega_c/\omega$ and $X = \omega_p^2/\omega^2$. Note that X is proportional to the electron density and Y is proportional to the magnetic field strength. In terms of X and Y the above equations can be rewritten as

$$R = 1 - \frac{X}{1 - Y}, \quad L = 1 - \frac{X}{1 + Y}, \quad S = 1 - \frac{X}{1 - Y^2}, \quad \text{and} \quad P = 1 - X. \quad (4.4.58)$$

Using these equations, it is easy to show that the bounding surfaces are given by

$R = 0,$	$X = 1 - Y$	cutoff,
$L = 0,$	$X = 1 + Y$	cutoff,
$P = 0,$	$X = 1$	cutoff,
$S = 0,$	$X = 1 - Y^2$	hybrid resonance,
$R = \infty,$	$Y = 1$	cyclotron resonance,
$L = \infty,$		no solution

The CMA diagram can then be constructed by plotting the bounding surfaces as a function of X and Y , as shown in Figure 4.36. The topological shapes of the index of refraction surfaces are summarized by the sketches in each bounded volume. Each sketch is labeled by the polarization (R , L) at $\theta = 0$, and the mode of propagation (O , X) at $\theta = \pi/2$.

In order to determine the accessible regions for each mode, the continuity of the index of refraction surfaces must be studied at each bounding surface. In some cases this is relatively easy, whereas in other cases it is quite difficult. For example, it is easy to see that the (L , O) mode is continuous throughout the region $X < 1$. On the other hand, the analysis of what happens to the extraordinary mode in the region around $X = 1$ and $Y = 1$ is quite complicated. In such complicated regions, the best approach is to investigate numerically the continuity across the boundary in question. Using such an approach it can be shown that there are only four continuously connected branches of the index of refraction. These four branches are indicated by the four shaded regions in Figures 4.37 and 4.38. Figure 4.37 is for the case $\omega_p > \omega_c$, and Figure 4.38 is for the case $\omega_p < \omega_c$.

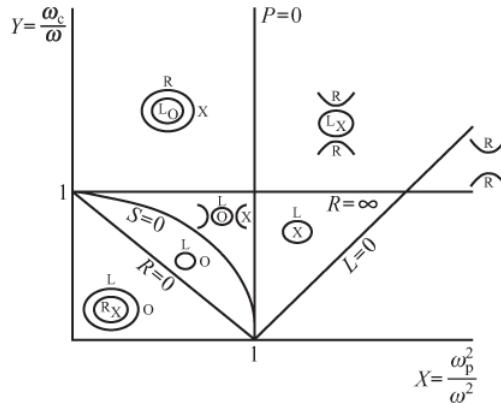


Fig. 4.36 The CMA diagram for a plasma of cold electrons and immobile ions.

The boundaries of the shaded regions are determined by the indices of refraction at $\theta = 0$ (solid line) and $\theta = \pi/2$ (dashed line). At intermediate wave normal angles, the index of refraction remains in the shaded regions. In both cases note that for $\theta = 0$ the (L , O) mode has a gap at $\omega = \omega_p$. Although the index of refraction at $\theta = 0$ is continuous across $\omega = \omega_p$, at any angle $\theta \neq 0$, the (L , O) mode splits into two branches at the plasma frequency. The high-frequency branch is the free space (L , O) mode, and the low-frequency branch is the Z-mode, named after a characteristic Z-shaped feature first identified in ground-based ionograms [Ratcliffe, 1962]. The Z-mode is bounded at low frequencies by the $\omega_{L=0}$ cut-off, and at high frequencies by the upper hybrid resonance ω_{UH} . The whistler mode is

bounded at high frequencies by either the electron cyclotron frequency, ω_c , or the electron plasma frequency, ω_p , whichever is smaller. The remaining mode, the free space R-X mode is bounded by $R = 0$. The regions of the CMA diagram that are accessible via the four continuously connected branches of the dispersion relation are summarized in Figure 4.39. The four branches can be grouped into modes that are accessible from free space ($X = 0, Y = 0$), and modes that are internally trapped (not accessible from free space). The modes accessible from free space are the right-hand polarized extraordinary mode (R, X) and the left-hand polarized ordinary mode (L, O). These two modes are of particular interest to radio astronomers because they are the only modes that can escape from an astrophysical radio source. In the ionosphere the cutoff frequencies of these two modes also give the low-frequency limit for ground-based radio astronomy. Typical plots of the cutoff frequencies, f_p and $f_{R=0}$, as a function of altitude in the Earth's ionosphere are shown in Figure 4.40. Left-hand and right-hand polarized radio waves arriving from a distant source are reflected from the top side of the ionosphere at f_p and $f_{R=0}$. Such waves can only reach the ground at frequencies above the maximum values for f_p and $f_{R=0}$, which are typically about 5 to 8 MHz. The same reflection process also occurs for waves incident on the ionosphere from below. Reflections from the bottom side of the ionosphere are important for long-range, over-the-horizon, radio communication.

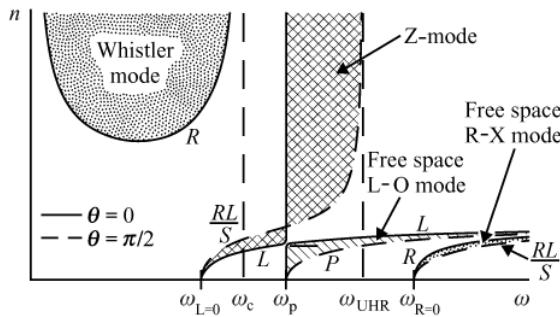


Fig. 4.37

The boundaries of the index of refraction for the four continuously connected branches of the dispersion relation for the case $\omega_p > \omega_c$.

The remaining two modes are the whistler and the Z-modes. These two modes are internally trapped and cannot reach free space ($X = 0$ and $Y = 0$). The CMA diagram shows that the whistler mode is bounded by $X > 1$ and $Y > 1$. The index of refraction of the whistler mode has the same topological form throughout the region where it exists. The Z-mode is more complicated and changes topological form at $X = 1$ and $Y = 1$. The Z-mode has a resonance cone within the region bounded by the lines $X = 1$, $Y = 1$, and $X = 1 - Y^2$. Elsewhere, the Z-mode has a quasi-ellipsoidal index of refraction surface. The boundary at $X = 1 - Y^2$ corresponds to the upper hybrid resonance. Along this boundary the Z-mode becomes almost purely electrostatic and the allowed range of wave normal angles shrinks to zero at $\theta = \pi/2$.

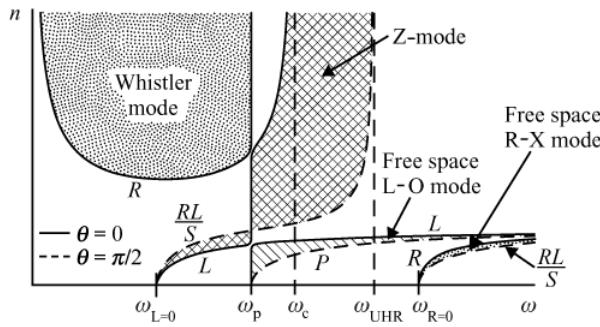


Fig. 4.38 The boundaries of the index of refraction for the four continuously connected branches of the dispersion relation for the case $\omega_p < \omega_c$.

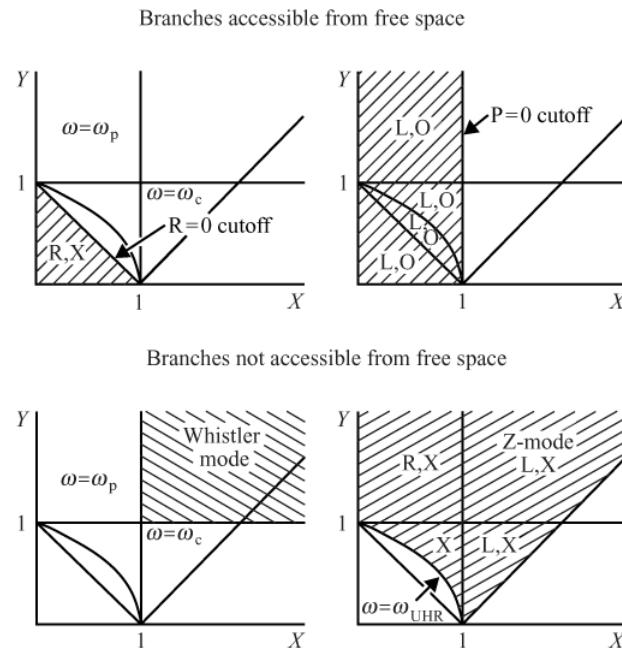


Fig. 4.39 The regions of the CMA diagram for the four continuously connected branches of the dispersion relation for a plasma of cold electrons and immobile ions.

4.4.4.2 The CMA diagram for electrons and one ion species

The inclusion of a single type of positive ion with finite mass does not appreciably alter the results presented in the previous section provided the wave frequency is above the lower hybrid resonance frequency. However, at frequencies below the lower hybrid resonance frequency, the effects of the finite ion mass become quite important.

Since charge neutrality demands that $n_{i0} = n_{e0}$, the CMA diagram for a two-component plasma still has only two parameters, just as when the ions were considered immobile. A convenient set of parameters for describing a two-component plasma are the normalized

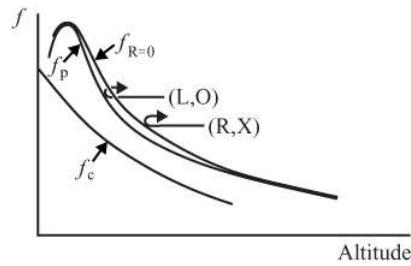


Fig. 4.40 A representative plot of the cutoff frequencies, f_p and $f_{R=0}$, as a function of altitude for the free space (R, X) and (L, O) modes.

electron cyclotron frequency $Y = |\omega_{ce}|/\omega$ and the normalized plasma frequency $X = (\omega_{pe}^2 + \omega_{pi}^2)/\omega^2$. The corresponding CMA diagram is shown in Figure 4.41. The two modes at large X and Y (in the upper right-hand corner) are the Alfvén modes described earlier. One of these modes is isotropic, and the other is highly anisotropic (i.e., it has a resonance cone). The isotropic Alfvén mode transforms smoothly and continuously into the whistler mode, and the anisotropic Alfvén mode transforms smoothly and continuously into the ion cyclotron mode. The ion cyclotron mode has a resonance at the ion cyclotron frequency and does not propagate above this frequency. At the ion cyclotron frequency, the ion cyclotron mode undergoes a destructive transition very similar to the destructive transition that occurs for the whistler mode at the electron cyclotron frequency.

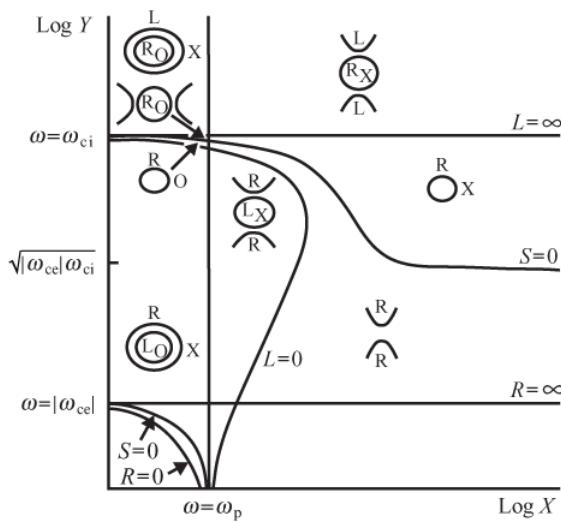


Fig. 4.41 The CMA diagram for a cold plasma of electrons and ions.

4.5 Ray paths in inhomogeneous plasmas

Next we return to the important question of refraction. Because the index of refraction changes in an inhomogeneous medium, the ray path is usually not a straight line. The bending of the ray path by inhomogeneous effects is called refraction. The problem of determining the ray path of a wave propagating in an inhomogeneous anisotropic medium is difficult. Meaningful solutions are possible only if the index of refraction varies slowly on a spatial scale comparable to the wavelength.

The problem of an electromagnetic wave propagating in a slowly varying medium can be solved by assuming a solution of the form $Ae^{i\phi}$, where the amplitude A is a slowly varying function of position. The function ϕ , which describes the wave front, is called the eikonal. Starting with Maxwell's equations and assuming that the dielectric properties are slowly varying functions of position, it is possible to derive an approximate set of equations ordered according to the magnitude of the inhomogeneity. These equations were first derived for isotropic optical media in the late nineteenth century and were extended to anisotropic media by Weinberg [1962]. Because of the complexity of the subject, it is our intention to give only a brief outline of the theory and focus on some simple applications.

For an inhomogeneous time dependent medium, Maxwell's equations can be written as a set of linear homogeneous partial differential equations of the form

$$\overset{\leftrightarrow}{D}(\nabla, \partial/\partial t, \mathbf{r}, t) \cdot \mathbf{f} = 0, \quad (4.5.1)$$

where \mathbf{f} is a generalized vector that includes all field components. In the eikonal approach it is assumed that $\overset{\leftrightarrow}{D}$ depends so weakly on r that all the rapid spatial dependence in \mathbf{f} can be represented by a common multiplicative factor of the form $e^{i\phi}$. With this assumption one can replace ∇ in $\overset{\leftrightarrow}{D}$ by $i\mathbf{k}$. If the temporal variations of the medium are slow one can further assume that \mathbf{f} depends on time through a factor $e^{-i\omega t}$, in which case $\partial/\partial t$ can be replaced by $-i\omega$. The system of linear differential equations given by equations (4.5.1) then becomes

$$\overset{\leftrightarrow}{D}(i\mathbf{k}, -i\omega, \mathbf{r}, t) \cdot \mathbf{f} = 0, \quad (4.5.2)$$

where $i\mathbf{k} = \nabla\phi$.

To solve the above equation for $\phi(r)$ one further assumes that \mathbf{k} must satisfy the dispersion relation for a homogeneous medium evaluated at \mathbf{r} , i.e., $D(\mathbf{k}, \omega, \mathbf{r}, t) = 0$. The dispersion relation by itself does not give \mathbf{k} . To find \mathbf{k} at all points along the ray path, one introduces a one-parameter family of trajectories in \mathbf{r}, \mathbf{k} space defined by

$$\mathbf{r} = \mathbf{r}(\tau) \text{ and } \mathbf{k} = \mathbf{k}(\tau). \quad (4.5.3)$$

A requirement is then imposed that the variation in the dispersion relation be zero along the actual ray path, i.e., $\delta D = 0$. Therefore, if the dispersion relation is satisfied at the initial point $\tau = 0$, it is satisfied at all points along the ray path. The variation in D is given by

$$\delta D = \nabla_k D \cdot \frac{d\mathbf{k}}{d\tau} d\tau + \frac{\partial D}{\partial \omega} \frac{d\omega}{d\tau} d\tau + \nabla D \cdot \frac{d\mathbf{r}}{d\tau} d\tau + \frac{\partial D}{\partial t} \frac{dt}{d\tau} d\tau = 0. \quad (4.5.4)$$

Equation (4.5.4) is satisfied if we require that \mathbf{r} , \mathbf{k} and ω satisfy the following system of equations:

$$\frac{d\mathbf{r}}{d\tau} = \nabla_k D, \quad (4.5.5)$$

$$\frac{d\mathbf{k}}{d\tau} = -\nabla D, \quad (4.5.6)$$

$$\frac{d\omega}{d\tau} = \frac{\partial D}{\partial t}, \quad (4.5.7)$$

and

$$\frac{dt}{d\tau} = -\frac{\partial D}{\partial \omega}. \quad (4.5.8)$$

Equation (4.5.8) can be regarded as the defining equation for the parameter τ .

To interpret this system of equations, note that equations (4.5.5) and (4.5.8) can be combined to give

$$\frac{d\mathbf{r}}{dt} = -\frac{\nabla_k D}{\partial D / \partial \omega} = \nabla_k \omega = \mathbf{v}_g. \quad (4.5.9)$$

Thus, the space-time trajectory is parallel to the group velocity, which by definition is the ray path. If we further identify D as the Hamiltonian, \mathbf{k} as the momentum, ω as the total energy, and τ as the time, the first three equations are seen to be identical to Hamilton's equations of classical mechanics. Thus, the ray path problem is formally identical to the problem of solving for the motion of a particle in a system with a Hamiltonian $D(\mathbf{k}, \omega, \mathbf{r}, t)$. The close correspondence between the motion of a wave packet and the motion of a particle should not be surprising, since the duality between waves and particles is a well-known feature of quantum mechanics. The correspondence between the ray path equations and Hamilton's equations also means that one can now make use of the many powerful techniques available in classical mechanics to solve particle trajectories.

Even though the ray path problem can in principle be solved like a problem in classical mechanics, in all but the simplest cases one must resort to numerical computations. One such case of special interest is a planar geometry in which the parameters of the medium depend on only one coordinate, z for example. The coordinates x and y are then cyclic coordinates, since $D(\mathbf{k}, \omega)$ does not depend on x and y . As in classical mechanics, it follows that the "conjugate momenta," k_x and k_y , are constants. By a suitable coordinate rotation we can then arrange for \mathbf{k} to be in the x, z plane, so that we need only concern ourselves with the condition $k_x = \text{constant}$. This condition is equivalent to Snell's law, which can be written

$$n_x = n \sin \theta = \text{constant.} \quad (4.5.10)$$

For a linearly polarized wave with the electric field in the y direction, it is straightforward to show that Maxwell's equations can be combined to give the following equation for E_y ,

$$\frac{d^2 E_y}{dz^2} + k_z^2(z) E_y = 0. \quad (4.5.11)$$

This equation is seen to be of the same form as the time varying harmonic oscillator discussed in Section ???. If k_z varies sufficiently slowly, an approximate solution is given by the Wentzel–Kramers–Brillouin (WKB) equation

$$E_y = \frac{A}{\sqrt{k_z}} \exp \left[\pm i \int^z k_z(z') dz' \right], \quad (4.5.12)$$

which is valid if the condition

$$k_z^2 \gg \left| \frac{3}{4} \left(\frac{1}{k_z} \frac{dk_z}{dz} \right)^2 - \frac{1}{2} \frac{1}{k_z} \frac{d^2 k_z}{dz^2} \right| \quad (4.5.13)$$

is satisfied. A similar solution can be obtained for E_x . Since k_x and k_y are determined by the initial conditions and are constant, k_z can be computed as a function of z from the dispersion relation. The integral in the WKB solution then gives the phase ϕ of the electric field at any arbitrary height.

The ray path for a planar geometry can be determined using a graphical construction introduced by Poeverlein [1950]. This construction makes use of the facts that the x component of the index of refraction vector is constant and that the group velocity is perpendicular to the index of refraction surface. To make a Poeverlein construction the index of refraction surface is drawn at a sequence of heights as shown in Figure 4.42. The wave is assumed to be approaching a cutoff, indicated by the horizontal hashed line labeled $n = 0$. The vertical coordinate is z and the horizontal coordinate is x . A vertical (dashed) line is drawn through the index of refraction surfaces at $n_x = n \sin \theta_0$, where θ_0 is the initial angle of incidence with respect to the vertical. The intersection of the dashed line and the index of refraction surfaces then gives the index of refraction vector at each height. Note that two solutions occur, one corresponding to an upgoing wave and the other to a downgoing wave. The group velocities can be determined for each of these solutions. Once the group velocity vectors are known the ray path can be determined by drawing a smooth curve parallel to the group velocity at each altitude. Starting with the upgoing solution the ray path proceeds upward and becomes horizontal at point 3. At this point, the group velocity is horizontal and the upgoing and downgoing solutions merge. This is a reflection point. The wave energy is then transferred to the downgoing solution, and the ray path continues downward. Note that the wave does not reach the cutoff surface where $n = 0$. This is because the wave is refracted away from the region where the index of refraction goes to zero. Thus, even though the index of refraction remains real and varies smoothly up to the

cutoff, the region above the reflection point is not accessible. Accessibility in this case is limited by refraction effects. The reflection point can be moved closer to the cutoff surface by decreasing θ_0 . For $\theta_0 = 0$, the wave can propagate all the way up to the cutoff. However, in the vicinity of the cutoff the WKB solution (4.5.11) fails because k_z goes to zero. In this region, a so-called full-wave solution is required to solve Maxwell's equations. Typically these solutions involve Airy functions. For a discussion of these techniques, see Stix [1992].

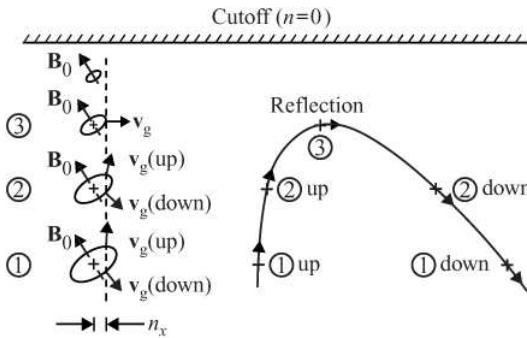


Fig. 4.42 A Poeeverlein construction illustrating the ray path of a wave approaching a cutoff in a horizontally stratified plasma.

If a cyclic coordinate does not exist, as often happens in two- or three-dimensional geometries, then the ray paths must be computed numerically using equations (4.5.5) through (4.5.8). Various numerical ray tracing codes are available for solving these equations, usually based on techniques developed by Haselgrove [1955].

Problems

- 4.1. Compute the Fourier transform of a Gaussian harmonic wave packet

$$f(x) = e^{-a_0^2 x^2/2} e^{ik_0 x}$$

where a_0 and k_0 are constants.

- 4.2. Show for the Gaussian harmonic wave packet that $\langle \Delta x^2 \rangle \langle \Delta k^2 \rangle = 1/4$.
- 4.3. A waveguide consists of two parallel conducting plates separated by a distance L . The dispersion relation of electromagnetic waves propagating in this waveguide is given by

$$\omega^2 = \omega_n^2 + c^2 k^2,$$

where ω_n is the cutoff frequency of the n th mode ($n = 1, 2, 3, \dots$, etc.) and $\omega_n = \pi c n / L$.

- (a) Show that the phase velocity is

$$v_p = \frac{\omega}{k} = \frac{c}{\sqrt{1 - \frac{\omega_p^2}{\omega^2}}}.$$

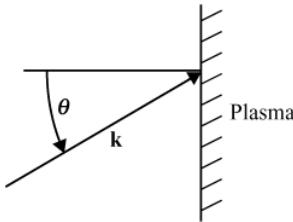
(b) Show that the group velocity is

$$v_g = \frac{\partial \omega}{\partial k} = c \sqrt{1 - \frac{\omega_p^2}{\omega^2}}.$$

(c) Plot v_p and v_g as a function of ω .

(d) Compare the above results to the propagation of electromagnetic waves in an unmagnetized plasma.

- 4.4. Waves are incident on a slab of plasma at an angle θ as shown below. The plasma has a uniform plasma frequency ω_p , and the index of refraction is given by $n = \sqrt{1 - \omega_p^2/\omega^2}$. Using Snell's law, find the critical angle θ_c for total internal reflection (within the free space region) as a function of frequency.



- 4.5. The effect of collisions can be included in the dispersion relation for waves in a cold plasma by adding a drag force $v_s m_s v_s$ to the equation of motion (v_s is called the collision frequency).

$$m_s \frac{d\mathbf{v}_s}{dt} + v_s m_s \mathbf{v}_s = e_s [\mathbf{E} + \mathbf{v}_s \times \mathbf{B}].$$

(a) Show that the effect of collisions can be obtained by making the substitution

$$m_s \rightarrow m_s \left(1 + \frac{iv_s}{\omega} \right)$$

in the collisionless dispersion relation.

(b) For transverse waves in a cold plasma, show that if $v_s \ll \omega$ and $\omega_p \ll \omega$ (the high frequency approximation) the real and imaginary parts of the wave number are approximately

$$k_r = \frac{\omega}{c} \left[1 - \frac{\omega_p^2}{2\omega^2} \right], \text{ and } k_i = \frac{1}{2c} \sum_s \frac{v_s \omega_{ps}^2}{\omega^2}.$$

(c) Find the dispersion relation for the longitudinal electron plasma oscillations including collisions. Briefly discuss the damping of these waves (what is the damping decrement and how does it depend on v_e).

- 4.6. In the derivation of the index of refraction for waves propagating in a cold plasma with a magnetic field, we defined the quantities $D = (1/2)(R - L)$ and $S = (1/2)(R + L)$. Prove the following identity:

$$S^2 - D^2 = RL.$$

- 4.7. Show that the determinant of the matrix

$$\begin{bmatrix} S - n^2 \cos^2 \theta & -iD & n^2 \sin \theta \cos \theta \\ iD & S - n^2 & 0 \\ n^2 \sin \theta \cos \theta & 0 & P - n^2 \sin^2 \theta \end{bmatrix}$$

can be written in the form

$$D(\mathbf{n}, \omega) = An^4 - Bn^2 + C$$

where

$$\begin{aligned} A &= S \sin^2 \theta + P \cos^2 \theta, \\ B &= RL \sin^2 \theta + PS(1 + \cos^2 \theta), \end{aligned}$$

and

$$C = RLP.$$

- 4.8. Show that $F^2 = B^2 - 4AC$ can be rewritten in the positive definite form

$$F^2 = (RL - PS)^2 \sin^4 \theta + 4P^2 D^2 \cos^2 \theta.$$

- 4.9. Show that in the limit $\omega \rightarrow 0$

$$R = L = S = 1 + \sum_s \frac{\omega_{ps}^2}{\omega_{cs}^2},$$

$$D = 0 \quad \text{and} \quad P = - \sum_s \frac{\omega_{ps}^2}{\omega^2}.$$

- 4.10. Assuming that the ions are infinitely massive, derive the equations for the following frequencies:

- (a) the right-hand cutoff $\omega_{R=0}$;
- (b) the left-hand cutoff $\omega_{L=0}$;
- (c) the upper hybrid resonance ω_{UHR} .

- 4.11. For a plasma consisting of electrons and an arbitrary number of ion species show that the lower hybrid resonance frequency is given by the following equation if $\omega_{ci} \ll \omega_{LHR} \ll \omega_{ce}$

$$\omega_{LHR}^2 = \frac{1}{\frac{1}{\omega_{ce}^2} + \frac{1}{\omega_{pe}^2}} \sum_i \left(\frac{m_e}{m_i} \right) \alpha_i, \quad \text{where } \alpha_i = \frac{n_i}{n_e}.$$

- 4.12. Assuming the frequency is sufficiently high that the ions do not move and that $\omega \ll \omega_p, \omega_c \ll \omega_p$, show that the index of refraction of the whistler mode is approximately

$$n^2 = \frac{\omega_p^2}{\omega(\omega_c \cos \theta - \omega)}.$$

Plot $n(\theta)$ for $\omega \ll \omega_c$, $\omega = (0.25)\omega_c$, and $\omega = (0.5)\omega_c$.

Hint: Use the approximations:

$$\begin{aligned} R &\simeq -\omega_p^2/\omega(\omega - \omega_c), \quad L \simeq -\omega_p^2/\omega(\omega + \omega_c), \\ S &\simeq -\omega_p^2/(\omega^2 - \omega_c^2), \quad D \simeq -\omega_p^2\omega_c/\omega(\omega^2 - \omega_c^2), \quad \text{and} \\ P &= -\omega_p^2/\omega^2. \end{aligned}$$

- 4.13. Using the above formula for the index of refraction, show that for parallel propagation ($\theta = 0$) the group velocity of the whistler mode is given by

$$v_g = 2c \frac{\omega^{1/2}(\omega_c - \omega)^{3/2}}{\omega_c \omega_p}.$$

- 4.14. For frequencies $\omega \ll \omega_c$, one can see from problem 4.12 that the index of refraction of the whistler mode is given by

$$n^2 = \frac{\omega_p^2}{\omega \omega_c \cos \theta}.$$

Show that the group velocity makes an angle ψ relative to the magnetic field that is given by

$$\tan \psi = \frac{\sin \theta \cos \theta}{1 + \cos^2 \theta}.$$

From the above equation, show that the group velocity makes a maximum angle of $\psi_{\max} = \tan^{-1}(l/\sqrt{8}) = 19^\circ 28'$ with respect to the magnetic field.

Hint: In the text, it is shown that the angle ψ is given by

$$\psi = \theta + \alpha, \quad \text{where} \quad \tan \alpha = -\frac{1}{n} \frac{\partial n}{\partial \theta}.$$

Evaluate $\partial \psi / \partial \theta = 0$ to find the maximum value of ψ .

- 4.15. A plasma consists of electrons and one ion species. Assuming that $\omega \ll \omega_{pe}$ and $\omega_{pi} \gg \omega_{ci}$, show that near resonance ($\omega \sim \omega_{ci}$) the index of refraction of the ion cyclotron mode is given by

$$n^2 \simeq n_A^2 \left(\frac{\omega_{ci}}{\omega_{ci} - \omega} \right) \frac{1 + \cos^2 \theta}{2 \cos^2 \theta},$$

where n_A is the Alfvén index of refraction.

Hint: Assume that P is sufficiently large that one can make the approximations $A \simeq P \cos^2 \theta$ and $B \simeq PS(1 + \cos^2 \theta)$. Near resonance one can assume that $L \gg R$ so

that $S \approx L/2$. Since $n_e = n_i$, one can also use the relationship $\omega_{pe}^2/|\omega_{ce}| = \omega_{pi}^2/\omega_{ci}$. Finally, note that $n_A^2 = \omega_{pi}^2/\omega_{ci}^2$ when $\omega_{pi}^2 \gg \omega_{ci}$.

- 4.16. Using the results of problem 4.15 above and the same procedure as in problem 4.14, show that the maximum group velocity angle of the ion cyclotron mode is $\psi_{\max} = 12.3^\circ$.

Hint: In this case

$$\tan \psi = \frac{\sin \theta \cos^3 \theta}{1 + \cos^4 \theta}.$$

Because of the complicated form of this equation, the best way to find ψ_{\max} is to use a computer to calculate ψ as a function of θ .

- 4.17 For a sufficiently large plasma density show that the Alfvén velocity $V_A = c/n_A$ is given to a good approximation by the following equation:

$$V_A = \sqrt{\frac{B^2}{\mu_0 \rho_m}},$$

where $\rho_m = \sum_s m_s n_s$ is the mass density.

- 4.18. A transverse wave on a wire of linear mass density λ and tension T propagates at a velocity $V = \sqrt{T/\lambda}$. Show that the Alfvén velocity can be obtained from the above formula if the magnetic field has a tension (per unit cross-sectional area) given by B^2/μ_0 .

- 4.19. Waves are sometimes classified as compressible or non-compressible according to whether there is or is not a density fluctuation associated with the wave. By examining the electric field eigenvector for the Alfvén modes and the particle equations of motion, show that the isotropic Alfvén mode, $n^2 = n_A^2$, is compressible, whereas the anisotropic Alfvén mode, $n^2 = n_A^2/\cos^2 \theta$, is incompressible.

- 4.20. A plasma consists of equal densities of electrons and positrons (equal masses, $m_e^- = m_e^+$) in a uniform, static magnetic field.

(a) Write down and simplify the equations for R, L, S, D , and P using only the parameters ω_p and ω_c , where $\omega_p = \omega_{pe}^- = \omega_{pe}^+$ and $\omega_c = |\omega_{ce}^-| = |\omega_{ce}^+|$.

(b) Using the parameters $X = \omega_p/\omega$ and $Y = \omega_c/\omega$, construct a CMA diagram showing all bounding surfaces.

(c) Show that the indices of refraction of the two modes are given by

$$n^2 = S \quad \text{and} \quad n^2 = \frac{PS}{S \sin^2 \theta + P \cos^2 \theta}.$$

(d) Sketch the shape of the index of refraction surfaces in each region of the CMA diagram.

(e) Using appropriate diagrams summarize the accessible region of the CMA diagram for each continuous branch of the index of refraction.

References

- Alfvén, H. 1942. Existence of electromagnetic-hydrodynamic waves. *Nature* **150**, 405.
- Allis, W. P. 1959. *Waves in Plasmas, Sherwood Conference on Controlled Fusion, April 27–28, 1959, Gatlinburg, TN*, p. 32.
- Barkhausen, H. 1919. Zwei mit Hilfe der neuen Verstärker entdeckte Erscheinungen. *Phys. Z.* **20**, 401–403.
- Benson, R. F. 1982. Stimulated plasma instability and nonlinear phenomena in the ionosphere. *Radio Sci.* **17** (6), 1637–1659.
- Brice, N. M., and Smith, R. L. 1965. Lower hybrid resonance emissions. *J. Geophys. Res.* **78** (1), 71–80.
- Clemmow, P. C., and Mullaly, R. F. 1955. Dependence of the refractive index in magneto-ionic theory on the direction of the wave normal. *Physics of the Ionosphere: Rep. Phys. Soc. Conf.* London Physical Society: Cavendish Laboratory, London, p. 340.
- Franklin, C. A., and Maclean, M. A. 1969. The design of swept-frequency topside sounders. *Proc. IEEE* **57**, 897–929.
- Gurnett, D. A., Shawhan, S. D., Brice, N. M., and Smith, R. L. 1965. Ion cyclotron whistlers. *J. Geophys. Res.* **70**, 1665–1688.
- Haselgrave, J. 1955. Ray theory and a new method for ray tracing. *Physics of the Ionosphere: Rep. Phys. Soc. Conf.* London Physical Society: Cavendish Laboratory, London, pp. 355–364.
- Helliwell, R. A. 1965. *Whistlers and Related Ionospheric Phenomena*. Stanford University Press: Stanford, CA, p. 103.
- Kaplan, W. 1952. *Advanced Calculus*. Addison-Wesley: Reading, MA, p. 99.
- Manchester, R. N., and Taylor, J. H. 1977. *Pulsars*. Freeman and Company: San Francisco, CA, p. 134.
- Poeverlein, H. 1950. Strahlwege von Radiowellen in der Ionosphäre. *Z. Angew. Phys.* (4), 152–160.
- Ratcliffe, J. A. 1962. *The Magneto-Ionic Theory and Its Application to the Ionosphere*. Cambridge University Press: London, p. 126.
- Stix, T. H. 1992. *Waves in Plasmas*. American Institute of Physics: New York, p. 343.
- Storey, L. R. O. 1953. An investigation of whistling atmospherics. *Philos. Trans. R. Soc. London, Ser. A* **246**, 113–141.
- Townsend, J. S. 1992. *A Modern Approach to Quantum Mechanics*. McGraw-Hill: New York, p. 79.
- Walsh, D., Haddock, T. F., and Schultz, H. F. 1964. Cosmic radio intensities at 1.225 and 2.0 MC measured up to an altitude of 1700 km. *Space Sci.* **IV**, 935–959.

- Weinberg, S., 1962. Eikonal methods in magnetohydrodynamics. *Phys. Rev.* **126**, 1899–1909.

Further reading

- Budden, K. G. *Radio Waves in the Ionosphere: The Mathematical Theory of the Reflection of Radio Waves from Stratified Ionised Layers*. Cambridge University Press: London, 1961.
- Goertz, C. K., and Strangeway, R. J. *Introduction to Space Physics*, ed. Kivelson, M. G., and Russell, C. T. Cambridge University Press: Cambridge, 1995 (reprinted in 1996, 1997).
- Parks, G. K. *Physics of Space Plasmas: An Introduction*. Addison-Wesley: Redwood City, CA, 1991 (reprinted in 1996, 1997), Chapter 9.
- Stix, T. H. *Waves in Plasmas*. American Institute of Physics: New York, 1992.
- Swanson, D. G. *Plasma Waves*. Academic Press, Harcourt Brace: San Diego, CA, 1989, Chapter 2.

Kinetic theory and the moment equations

In the previous chapter we considered the idealized case of a cold plasma in which there were no random thermal motions. In this chapter we introduce a more general framework for analyzing plasmas in which thermal motions must be considered. For a system with a large number of particles it is neither possible nor desirable to determine the motion of each and every particle. Instead, we will use a statistical approach to compute the average motion of a large number of particles. This approach is called kinetic theory. It is not our intention in this chapter to present an exhaustive treatment of plasma kinetic theory. Instead we will introduce the basic concepts of kinetic theory and derive a system of equations known as the moment equations. These equations will then be used to analyze various simple applications that are of interest.

5.1 The distribution function

To carry out a statistical description of a plasma, it is convenient to introduce a six-dimensional space, called *phase space*, that consists of the position coordinates x, y , and z , and the velocity coordinates v_x, v_y , and v_z . At any given time the dynamical state of a particle can be represented by a point in phase space. For a system of many particles, the dynamical state of the entire system can then be represented by a collection of points in phase space, with one point for each particle. If the number of particles is very large, it is meaningful to define the average number density of particles in a small volume element of phase space. This density is called the distribution function $f(\mathbf{r}, \mathbf{v}, t)$ and is defined such that

$$dN = f(\mathbf{r}, \mathbf{v}, t) d^3x d^3v, \quad (5.1.1)$$

is the number of particles in the phase-space volume element $d^3x d^3v = dx dy dz dv_x dv_y dv_z$ at time t . We shall assume that the number of particles is sufficiently large that $f(\mathbf{r}, \mathbf{v}, t)$ can be regarded as a continuous function of \mathbf{r} and \mathbf{v} . The total number of particles, N , in the system is obtained by integrating $f(\mathbf{r}, \mathbf{v}, t)$ over all of phase space

$$N = \int_{-\infty}^{\infty} f(\mathbf{r}, \mathbf{v}, t) d^3x d^3v. \quad (5.1.2)$$

Having defined the distribution function we can then use the standard rules of statistics for computing various macroscopic averages. The average (expectation) value of any dynamical quantity $g(\mathbf{r}, \mathbf{v})$ in a region R of phase space is given by

$$\langle g(\mathbf{r}, \mathbf{v}) \rangle = \frac{1}{N} \int_R g(\mathbf{r}, \mathbf{v}) f(\mathbf{r}, \mathbf{v}, t) d^3x d^3v, \quad (5.1.3)$$

where the angle brackets, $\langle \rangle$, mean “average.”

Since several species may be present in a plasma it is necessary to define a distribution function $f_s(\mathbf{r}, \mathbf{v}, t)$ for each particle species s in the plasma. By making suitable choices for $g(\mathbf{r}, \mathbf{v})$ we can then compute various types of averages. Some of the most commonly used averages are the following:

$$\text{number density, } n_s = \int_V f_s(\mathbf{r}, \mathbf{v}, t) d^3v, \quad (5.1.4)$$

$$\text{average velocity, } \mathbf{U}_s = \frac{1}{n_s} \int_V \mathbf{v} f_s(\mathbf{r}, \mathbf{v}, t) d^3v, \quad (5.1.5)$$

$$\text{kinetic energy density, } W_s = \int_V \frac{1}{2} m_s v^2 f_s(\mathbf{r}, \mathbf{v}, t) d^3v, \quad (5.1.6)$$

$$\text{pressure tensor, } \overset{\leftrightarrow}{\mathbf{P}}_s = \int_V m_s (\mathbf{v} - \mathbf{U}_s)(\mathbf{v} - \mathbf{U}_s) f_s(\mathbf{r}, \mathbf{v}, t) d^3v, \quad (5.1.7)$$

where the integrations are over velocity space, V . The number density, n_s , is the average number of particles of type s per unit volume. The velocity, \mathbf{U}_s , is the average velocity of particles of type s . The kinetic energy density, W_s , is the average kinetic energy, $(1/2)m_s v^2$, of particles of type s per unit volume. In index notation, the pressure tensor, $\overset{\leftrightarrow}{\mathbf{P}}_s = [P_{ij}]$, is the average rate at which momentum is transported in the i direction across surface j in a frame of reference moving at the average velocity, \mathbf{U}_s . The quantity $(\mathbf{v} - \mathbf{U}_s)(\mathbf{v} - \mathbf{U}_s)$ in equation (5.1.7) is defined by the following matrix,

$$\begin{aligned} & (\mathbf{v} - \mathbf{U}_s)(\mathbf{v} - \mathbf{U}_s) \\ &= \begin{bmatrix} (v_x - U_{sx})(v_x - U_{sx}), (v_x - U_{sx})(v_y - U_{sy}), (v_x - U_{sx})(v_z - U_{sz}) \\ (v_y - U_{sy})(v_x - U_{sx}), (v_y - U_{sy})(v_y - U_{sy}), (v_y - U_{sy})(v_z - U_{sz}) \\ (v_z - U_{sz})(v_x - U_{sx}), (v_z - U_{sz})(v_y - U_{sy}), (v_z - U_{sz})(v_z - U_{sz}) \end{bmatrix}. \end{aligned} \quad (5.1.8)$$

The average number density, n_s , and the average velocity, \mathbf{U}_s , can be used to compute the charge density, ρ_q , and the mass density, ρ_m , using the equations

$$\rho_q = \sum_s e_s n_s, \quad (5.1.9)$$

and

$$\rho_m = \sum_s m_s n_s. \quad (5.1.10)$$

Similarly, the current density is given by

$$\mathbf{J} = \sum_s e_s n_s \mathbf{U}_s. \quad (5.1.11)$$

To illustrate the methods used to compute the various averages defined above, it is instructive to work out the integrals involved for certain specific choices of the distribution function. For example, consider a distribution of electrons all moving at a velocity $\mathbf{v} = \mathbf{U}_0$ and of ions all at rest, $\mathbf{v} = 0$. The velocity distribution functions in this case can be written

$$f_e = n_0 \delta(\mathbf{v} - \mathbf{U}_0) \quad \text{and} \quad f_i = n_0 \delta(\mathbf{v}), \quad (5.1.12)$$

where δ is the Dirac delta function. For this simple case, it is easy to show using equation (5.1.4) that $n_e = n_i = n_0$. The charge density is then $\rho_q = 0$. The average velocity of the electrons, also sometimes called the drift velocity, is $\mathbf{U}_e = \mathbf{U}_0$, and the average velocity of the ions is $\mathbf{U}_i = 0$. Since the electrons are drifting with respect to the ions, a current is present. The corresponding current density is given by $\mathbf{J} = -en_0\mathbf{U}_0$. Since there is no spread in velocity with respect to the average velocity, the pressure tensors $\overset{\leftrightarrow}{\mathbf{P}}_e$ and $\overset{\leftrightarrow}{\mathbf{P}}_i$ are zero.

Next, consider the more complicated case of a Maxwellian velocity distribution function given by

$$f = n_0 \left(\frac{m}{2\pi\kappa T} \right)^{3/2} \exp \left[-\frac{m(\mathbf{v} - \mathbf{U}_0)^2}{2\kappa T} \right], \quad (5.1.13)$$

where for simplicity the subscript s has been omitted. This distribution function has a peak at velocity $\mathbf{v} = \mathbf{U}_0$ and is spherically symmetric around this velocity. It is easy to show that the average velocity, defined by equation (5.1.5), is simply \mathbf{U}_0 . The integrals in the pressure tensor can also be easily computed and are

$$p_{ij} = n_0 k T \quad \text{for } i = j, \quad (5.1.14)$$

and

$$p_{ij} = 0 \quad \text{for } i \neq j, \quad (5.1.15)$$

The pressure tensor is diagonal, with equal elements, which indicates an isotropic pressure. The pressure is also directly proportional to the number density, n_0 , and the temperature, T , which is simply the ideal gas law, $P = n_0 k T$.

If the distribution function is anisotropic, the diagonal elements of the pressure tensor are in general not the same. An example of an anisotropic distribution function is given by the bi-Maxwellian

$$f = n_0 \left(\frac{m}{2\pi\kappa T_\perp} \right) \left(\frac{m}{2\pi\kappa T_\parallel} \right)^{1/2} \exp \left[-\frac{mv_\perp^2}{2\kappa T_\perp} \right] \exp \left[-\frac{mv_\parallel^2}{2\kappa T_\parallel} \right], \quad (5.1.16)$$

which is often used in a magnetized plasma. The symbols \perp and \parallel refer to directions perpendicular and parallel to the magnetic field (i.e., cylindrical coordinates). If we choose the coordinate system such that the magnetic field is aligned with the z axis, then the pressure tensor, defined by equation (5.1.7), is diagonal and has the simple form

$$\overset{\leftrightarrow}{\mathbf{P}} = \begin{bmatrix} p_{\perp} & 0 & 0 \\ 0 & p_{\perp} & 0 \\ 0 & 0 & p_{\parallel} \end{bmatrix}, \quad (5.1.17)$$

where $P_{\perp} = n_0 \kappa T_{\perp}$ and $P_{\parallel} = n_0 \kappa T_{\parallel}$. In the absence of collisions the parallel and perpendicular pressures are decoupled. Later we will discuss an equation of state, called the Chew, Goldberger, and Low equation of state, that describes how the perpendicular and parallel pressures respond to dynamical changes.

5.2 The Boltzmann and Vlasov equations

The kinetic equation that governs the evolution of the distribution function in phase space is a differential equation called the *Boltzmann equation*. To derive this equation, one starts by assuming that the forces acting on the particles can be classified into two types: long-range and short-range. Forces that are the same for all particles in a given phase-space volume element are long-range forces, and forces that are not the same are short-range forces. Long-range forces originate from the collective effects of a large number of particles acting over relatively long distances, and short-range forces originate from a small number of particles (usually only two) acting over relatively short distances. Short-range forces correspond to what are commonly called collisions. Collisions produce large, nearly instantaneous, changes in the velocities of the interacting particles.

To derive the Boltzmann equation, we first assume that there are no short-range forces (i.e., no collisions). All particles in an infinitesimal volume of phase space, $d^3x d^3v$, then experience the same long-range force \mathbf{F} . Since the force on a representative particle in phase space is given by Newton's law

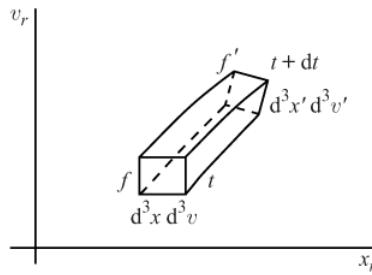


Fig. 5.1 The evolution of a volume element $d^3x d^3v$ in phase space from time t to $t + dt$.

$$\mathbf{F} = m \frac{d\mathbf{v}}{dt}, \quad (5.2.1)$$

it follows that the particle trajectories can never intersect. If the trajectories did intersect, two different values of $d\mathbf{v}/dt$ would exist for the same \mathbf{F} and \mathbf{v} , which cannot occur. Therefore, if we follow the evolution of a volume element that moves with the particles, then all of the particles within the volume element $d^3x \, d^3v$ at time t map into the corresponding volume element $d^3x' \, d^3v'$ at time $t' = t + dt$. The evolution of a typical phase-space volume element is illustrated in Figure 5.1. Since the number of particles in the phase-space volume element is conserved, we can write

$$f(\mathbf{r}, \mathbf{v}, t) \, d^3x \, d^3v = f'(\mathbf{r}', \mathbf{v}', t') \, d^3x' \, d^3v', \quad (5.2.2)$$

where f is the distribution function at time t , and f' is the distribution function at time t' . The equations of motion for the particles in phase space are given by

$$x'_r = x_r + v_r dt, \quad \text{and} \quad v'_r = v_r + \frac{F_r}{m} dt, \quad (5.2.3)$$

where we use index notation $r = (1, 2, 3)$ to represent $x_r = (x, y, z)$ and $v_r = (v_x, v_y, v_z)$. These equations can be viewed as simply a transformation from phase-space coordinates (\mathbf{r}, \mathbf{v}) to coordinates $(\mathbf{r}', \mathbf{v}')$. The infinitesimal volume elements involved in this transformation are determined by

$$d^3x' \, d^3v' = J \, d^3x \, d^3v, \quad (5.2.4)$$

where J is the Jacobian of the transformation [Kaplan, 1952]. Since there are six transformation equations the Jacobian is given by the 6×6 determinant

$$J = \text{Det} \begin{bmatrix} \frac{\partial x'}{\partial x} & \frac{\partial x'}{\partial y} & \dots \\ \frac{\partial y'}{\partial x} & \frac{\partial y'}{\partial y} & \dots \\ \vdots & \vdots & \ddots \\ \frac{\partial v'_x}{\partial x} & \frac{\partial v'_x}{\partial y} & \dots \\ \vdots & \vdots & \ddots \\ \frac{\partial v'_z}{\partial x} & \frac{\partial v'_z}{\partial y} & \dots \end{bmatrix}. \quad (5.2.5)$$

After computing the appropriate derivatives and sorting out the lowest order terms in dt , it is easy to show that equation (5.2.4) can be written in the form

$$d^3x' \, d^3v' = \left[1 + \sum_r \frac{\partial v_r}{\partial x_r} dt + \sum_r \frac{\partial}{\partial v_r} \left(\frac{F_r}{m} \right) dt \right] d^3x \, d^3v, \quad (5.2.6)$$

where the summation extends over $r = 1, 2$, and 3 (i.e., x, y , and z). To complete the derivation, the function f' in equation (5.2.2) is expanded in a Taylor series using the chain rule, with $v_r = dx_r/dt$ and $dv_r/dt = F_r/m$, which gives

$$f' = f + \frac{\partial f}{\partial t} dt + \sum_r v_r \frac{\partial f}{\partial x_r} dt + \sum_r \left(\frac{F_r}{m} \right) \frac{\partial f}{\partial v_r} dt. \quad (5.2.7)$$

Substituting $d^3x' d^3v'$ from equation (5.2.6) and f' from equation (5.2.7) into equation (5.2.2) then gives

$$f = f + \frac{\partial f}{\partial t} dt + \sum_r v_r \frac{\partial f}{\partial x_r} dt + \sum_r \left(\frac{F_r}{m} \right) \frac{\partial f}{\partial v_r} dt + \sum_r f \frac{\partial v_r}{\partial x_r} dt + \sum_r f \frac{\partial}{\partial v_r} \left(\frac{F_r}{m} \right) dt, \quad (5.2.8)$$

which simplifies to the first-order partial differential equation

$$\frac{\partial f}{\partial t} + \sum_r \frac{\partial}{\partial x_r} (v_r f) + \sum_r \frac{\partial}{\partial v_r} \left(\frac{F_r}{m} f \right) = 0. \quad (5.2.9)$$

Comparing the above equation with the familiar three-dimensional mass continuity equation of fluid mechanics,

$$\frac{\partial \rho_m}{\partial t} + \sum_r \frac{\partial}{\partial x_r} (\nu_r \rho_m) = 0, \quad (5.2.10)$$

we see that it is simply a continuity equation in six-dimensional phase space. Equation (5.2.9) is valid for both relativistic and non-relativistic velocities.

Two simplifications can be made to equation (5.2.9). First, note that since v_r and x_r are independent variables, it follows that $\partial v_r / \partial x_r$ is zero. The independence of v_r and x_r may not be immediately obvious, since in particle mechanics, $\mathbf{v}(t)$ is obtained as a time derivative of $\mathbf{r}(t)$ and hence depends on $\mathbf{r}(t)$. However, the kinetic description is not based on following an individual particle in phase space. Instead, it is based on the average number of particles in a volume element of phase space at any given time. Since the spatial co-ordinates of the volume element can be chosen independent of the velocity, x_r and v_r are independent variables. Second, note that if we restrict our consideration to non-relativistic velocities, for which the mass m is independent of velocity, the term $(\partial/\partial v_r)(F_r/m)$ is zero. This result can be proven for electromagnetic forces by simply working out the appropriate derivatives. For $r = x$, it is easily seen that

$$\frac{\partial}{\partial v_x} \left(\frac{F_x}{m} \right) = \frac{q}{m} \frac{\partial}{\partial v_x} [E_x + v_y B_z - v_z B_y] = 0. \quad (5.2.11)$$

The same result holds for $r = y$ and $r = z$. These two simplifications imply that the Jacobian is $J = 1$. The phase-space flow is then volume preserving, that is, $d^3x' d^3v' = d^3x d^3v$. With these simplifications equation (5.2.9) reduces to

$$\frac{\partial f}{\partial t} + \sum_r v_r \frac{\partial f}{\partial x_r} + \sum_r \frac{F_r}{m} \frac{\partial f}{\partial v_r} = 0, \quad (5.2.12)$$

or in vector notation

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{v}} f + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f = 0, \quad (5.2.13)$$

where $\nabla_{\mathbf{v}}$ is the gradient operator in velocity space, $\nabla_{\mathbf{v}} = \hat{\mathbf{x}}(\partial/\partial v_x) + \hat{\mathbf{y}}(\partial/\partial v_y) + \hat{\mathbf{z}}(\partial/\partial v_z)$. This result does not hold for relativistic velocities, since the mass in equation (5.2.11) depends on the velocity, i.e., $m = m_0/(1 - v^2/c^2)^{1/2}$, where m_0 is the rest mass.

Next, consider the effect of short-range forces (i.e., collisions). Because short-range forces are not the same for all particles in a volume element, collisions act to scatter particles into and out of the phase-space volume element. As mentioned earlier, the main effect of short-range interactions is to introduce a nearly discontinuous change in the particle velocity. Some typical trajectories in phase space are shown in Figure 5.2. Changes to the distribution function due to collisions as the phase-space volume element evolves from time t to time $t + dt$ are taken into account by adding a collision operator, $\delta_c f/\delta t$, not yet specified, to the right-hand side of equation (5.2.13), i.e.,

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{v}} f + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f = \frac{\delta_c f}{\delta t}. \quad (5.2.14)$$

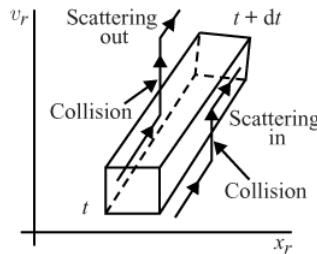


Fig. 5.2 Short-range forces (i.e., collisions) scatter particles out of and into the phase-space volume element as it evolves.

This equation is called the Boltzmann equation, after Boltzmann [1896–1898] who first used this equation to analyze particle transport phenomena in ordinary gases. The left-hand side of the Boltzmann equation describes the response of the particles to long-range collective forces, and the right-hand side describes the response to short-range collisional forces. As discussed earlier, if the number of electrons per Debye cube is large, $n\lambda_D^3 \gg 1$, then collective forces are much more important than collisional forces. Under these conditions the collision term can be ignored, i.e., $\delta_c f/\delta t = 0$. The resulting equation, with the force given by the Lorentz force, $\mathbf{F} = q[\mathbf{E} + \mathbf{v} \times \mathbf{B}]$, is

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{v}} f + \frac{q}{m} [\mathbf{E} + \mathbf{v} \times \mathbf{B}] \cdot \nabla_{\mathbf{v}} f = 0. \quad (5.2.15)$$

This equation is called the Vlasov equation, after Vlasov [1945] who first applied this equation to the study of plasmas.

5.3 Solutions based on constants of the motion

If a system has one or more known constants of the motion, then it is possible to write down immediately a general form of the solution to the Vlasov equation. These solutions are based on the fact that the distribution function is constant along a representative trajectory in phase space. To prove this result, compute the total derivative of f along a representative trajectory, $\mathbf{r}(t)$ and $\mathbf{v}(t)$, using the chain rule

$$df_s = \frac{\partial f_s}{\partial t} dt + \sum_r \frac{\partial f_s}{\partial x_r} \frac{dx_r}{dt} dt + \sum_r \frac{\partial f_s}{\partial v_r} \frac{dv_r}{dt} dt. \quad (5.3.1)$$

After recognizing that $dx_r/dt = v_r$ and $dv_r/dt = F_r/m$, and dividing by dt one obtains

$$\frac{df_s}{dt} = \frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} f_s = 0, \quad (5.3.2)$$

which is zero, since the right-hand side is just the Vlasov equation. Since $df_s/dt = 0$, it follows that f_s is constant along the trajectory. The possible trajectories are constrained by the constants of the motion, $C_m(\mathbf{r}, \mathbf{v})$. The constants of the motion can be obtained from the adiabatic invariants, as in Section ??, or from Hamilton's equations, as in Section ???. Any function of the constants of the motion

$$f_s[C_1(\mathbf{r}, \mathbf{v}), C_2(\mathbf{r}, \mathbf{v}), \dots] \quad (5.3.3)$$

is then a solution of the Vlasov equation.

As an example of this procedure, consider the motion of a particle of mass m and charge q in a time stationary electrostatic potential $\Phi(\mathbf{r})$. In this system the total energy

$$W = \frac{1}{2}mv^2 + q\Phi(\mathbf{r}) \quad (5.3.4)$$

is a constant of the motion. Any function $f_s[W(\mathbf{r}, \mathbf{v})]$ is then a solution of the Vlasov equation. The exact form of $f_s[W]$ is determined by the boundary conditions. For example, if $f(\mathbf{r}, \mathbf{v})$ is known to be a Maxwellian of number density n_0 and temperature T_0 at infinity, where $\Phi(\infty) = 0$, then the general solution at any arbitrary point is

$$f(\mathbf{r}, \mathbf{v}) = n_0 \left(\frac{m}{2\pi k T} \right)^{3/2} e^{-[\frac{1}{2}mv^2 + q\Phi(\mathbf{r})]/kT_0}. \quad (5.3.5)$$

This distribution function can be recognized as the function used in equation (??) to determine the potential in the plasma sheath around a point charge.

When constructing solutions of the Vlasov equation using constants of the motion, one must be careful about the boundary conditions. If the electrostatic potential is not monotonic, such as the example illustrated in Figure 5.3, then certain regions of phase space (x, v) may not be accessible from the boundary point x_0 . An example of such an inaccessible region is shown by the cross-hatching. Since the trajectories in the cross-hatched

region do not go through the point x_0 , the boundary conditions on f at x_0 do not constrain the distribution function in the inaccessible region. Some other boundary condition is then necessary to specify the distribution function. If no source is present in the inaccessible region, then the distribution function is zero in this region. In practice, collisions and various wave-particle interactions often scatter particles into the inaccessible region. An exact determination of the distribution function in such cases is usually quite difficult and must take into account these scattering effects.

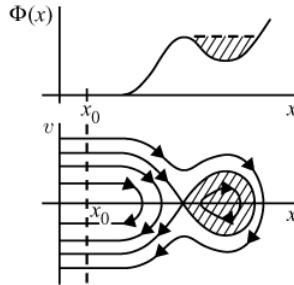


Fig. 5.3 The bottom (x, v) diagram shows the phase-space trajectories of the particles moving in the one-dimensional potential $\Phi(x)$ shown in the top diagram.

5.4 The moment equations

Often in analyzing problems in plasma physics, we are more interested in certain macroscopic averages such as the density and pressure rather than the detailed form of the distribution function. It is then sometimes convenient to use a system of equations called the moment equations. The moment equations are obtained by multiplying the Boltzmann equation by powers of the velocity and integrating over all of velocity space. This procedure generates one moment equation for each power of the velocity. We illustrate the procedure involved by calculating the first three moments of the Boltzmann equation.

5.4.1 Zeroth moment

To obtain the zeroth moment, we multiply the Boltzmann equation (5.2.14) by $v^0 = 1$ and integrate over all of velocity space, V , which gives

$$\int_V \frac{\partial f_s}{\partial t} d^3v + \int_V \mathbf{v} \cdot \nabla f_s d^3v + \frac{1}{m_s} \int_V \mathbf{F} \cdot \nabla_{\mathbf{v}} f_s d^3v = \int_V \frac{\delta_c f_s}{\delta t} d^3v. \quad (5.4.1)$$

The first integral on the left of this equation can be written

$$\int_V \frac{\partial f_s}{\partial t} d^3v = \frac{\partial}{\partial t} \int_V f_s d^3v = \frac{\partial n_s}{\partial t}, \quad (5.4.2)$$

where n_s is the average number density. Similarly, the second integral on the left can be written

$$\int_V \mathbf{v} \cdot \nabla f_s d^3v = \nabla \cdot \int_V \mathbf{v} f_s d^3v = \nabla \cdot (n_s \mathbf{U}_s), \quad (5.4.3)$$

where \mathbf{U}_s is the average velocity. Noting that $\nabla_{\mathbf{v}} \cdot (\mathbf{v} \times \mathbf{B}) f_s = (\mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} f_s$, the third integral on the left can be rewritten

$$\int_V \mathbf{F} \cdot \nabla_{\mathbf{v}} f_s d^3v = \int_V \nabla_{\mathbf{v}} \cdot (\mathbf{F} f_s) d^3v = \int_S (\mathbf{F} f_s) \cdot d\mathbf{s}_{\mathbf{v}} = 0, \quad (5.4.4)$$

where we have used Gauss' theorem to convert the volume integral to a surface integral at infinity. The surface integral is assumed to vanish if $f(\mathbf{v})$ decays sufficiently rapidly at infinite velocity. The integral on the right side of equation (5.4.1) involving the collision term is zero,

$$\int_V \frac{\delta_c f_s}{\delta t} d^3v = \frac{\delta_c n_s}{\delta t} = 0, \quad (5.4.5)$$

since collisions do not change the density. Combining equations (5.4.2)-(5.4.5), the zeroth moment of the Boltzmann equation (5.4.1) yields

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{U}_s) = 0. \quad (5.4.6)$$

This equation is simply the continuity equation for the s th species. Multiplying the above equation by the charge e_s and summing over all species, one obtains the charge continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0, \quad (5.4.7)$$

which is a fundamental equation of electricity and magnetism.

5.4.2 First moment

The first moment is obtained by multiplying the Boltzmann equation (5.2.14) by \mathbf{v} and integrating over all of velocity space. For reasons that will soon become apparent, it is more convenient to multiply by the momentum $m_s \mathbf{v}$, which gives

$$\begin{aligned} m_s \int_V \mathbf{v} \frac{\partial f_s}{\partial t} d^3v + m_s \int_V \mathbf{v} (\mathbf{v} \cdot \nabla f_s) d^3v + \int_V \mathbf{v} (\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3v \\ = m_s \int_V \mathbf{v} \frac{\delta_c f_s}{\delta t} d^3v. \end{aligned} \quad (5.4.8)$$

The first integral on the left of the above equation can be written

$$m_s \int_V \mathbf{v} \frac{\partial f_s}{\partial t} d^3v = \frac{\partial}{\partial t} \left(m_s \int_V \mathbf{v} f_s d^3v \right) = \frac{\partial}{\partial t} (m_s n_s \mathbf{U}_s). \quad (5.4.9)$$

The second integral is more complicated and can be written

$$m_s \int_V \mathbf{v} (\mathbf{v} \cdot \nabla f_s) d^3v = \nabla \cdot \left[m_s \int_V \mathbf{v} \mathbf{v} f_s d^3v \right]. \quad (5.4.10)$$

It is convenient to rewrite the $\mathbf{v}\mathbf{v}$ tensor by making use of the following identity,

$$\mathbf{v}\mathbf{v} = (\mathbf{v} - \mathbf{U}_s)(\mathbf{v} - \mathbf{U}_s) + \mathbf{U}_s \mathbf{v} + \mathbf{v} \mathbf{U}_s - \mathbf{U}_s \mathbf{U}_s. \quad (5.4.11)$$

The integral in equation (5.4.10) then simplifies to

$$m_s \int_V \mathbf{v}\mathbf{v} f_s d^3v = m_s \int_V (\mathbf{v} - \mathbf{U}_s)(\mathbf{v} - \mathbf{U}_s) f_s d^3v + m_s n_s \mathbf{U}_s \mathbf{U}_s. \quad (5.4.12)$$

The first term on the right in the above equation is just the pressure tensor, defined by equation (5.1.7). Equation (5.4.10) can then be written

$$m_s \int_V \mathbf{v} (\mathbf{v} \cdot \nabla f_s) d^3v = \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_s + \nabla \cdot (m_s n_s \mathbf{U}_s \mathbf{U}_s). \quad (5.4.13)$$

The third integral on the left of equation (5.4.8) is of the form

$$\begin{aligned} \int_V \mathbf{v} (\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3v &= \int_V (v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} + v_z \hat{\mathbf{z}}) \left[\frac{\partial}{\partial v_x} (F_x f_s) \right. \\ &\quad \left. + \frac{\partial}{\partial v_y} (F_y f_s) + \frac{\partial}{\partial v_z} (F_z f_s) \right] d^3v, \end{aligned} \quad (5.4.14)$$

where we have used the fact that for the Lorentz force

$$\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s = \nabla_{\mathbf{v}} \cdot (\mathbf{F} f_s). \quad (5.4.15)$$

Of the nine integrals that must be computed, six are of the type

$$\int_v v_y \frac{\partial}{\partial v_z} (F_z f_s) d^3v = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dv_x dv_y v_y \int_{-\infty}^{\infty} \frac{\partial}{\partial v_z} (F_z f_s) dv_z, \quad (5.4.16)$$

which can be integrated directly to give

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial v_z} (F_z f_s) dv_z = [F_z f_s]_{-\infty}^{\infty}. \quad (5.4.17)$$

If we assume that f_s goes to zero sufficiently rapidly at infinite velocity, these integrals are all zero. The remaining three integrals are of the type

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dv_x dv_y \int_{-\infty}^{\infty} dv_z v_z \frac{\partial}{\partial v_z} (F_z f_s), \quad (5.4.18)$$

which, after integrating by parts once, gives

$$\int_{-\infty}^{\infty} dv_z v_z \frac{\partial}{\partial v_z} (F_z f_s) = [v_z F_z f_s]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} F_z f_s dv_z. \quad (5.4.19)$$

If we again assume that f_s goes to zero sufficiently rapidly at infinite velocity, the first term on the right is zero. Combining all of these integrals, equation (5.4.14) can be written

$$\int_V \mathbf{v}(\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3v = - \int_V \mathbf{F} f_s d^3v. \quad (5.4.20)$$

After substituting the Lorentz force for \mathbf{F} and integrating, the above equation yields the simple result

$$\int_V \mathbf{v}(\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3v = -e_s n_s (\mathbf{E} + \mathbf{U}_x \times \mathbf{B}). \quad (5.4.21)$$

Finally, the first moment of the collision term on the right of equation (5.4.8) can be written

$$m_s \int_V \mathbf{v} \frac{\delta_c f_s}{\delta t} d^3v = \frac{\delta_c \mathbf{P}_s}{\delta t}, \quad (5.4.22)$$

where $\delta_c \mathbf{P}_s / \delta t$ is the average rate of change of the momentum per unit volume due to collisions. This term, which represents the drag force due to collisions, will be discussed in greater detail in Section 5.6 and again in Chapter ??.

Using equations (5.4.9), (5.4.13), (5.4.21), and (5.4.22), equation (5.4.8) becomes

$$\frac{\partial}{\partial t} (m_s n_s \mathbf{U}_s) + \nabla \cdot (m_s n_s \mathbf{U}_s \mathbf{U}_s) = n_s e_s (\mathbf{E} + \mathbf{U}_s \times \mathbf{B}) - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_s + \frac{\delta_c \mathbf{P}_s}{\delta t}. \quad (5.4.23)$$

The two terms on the left of this equation can be written in a more useful form by expanding the time derivative and using the identity $\nabla \cdot (\mathbf{AB}) = (\nabla \cdot \mathbf{A})\mathbf{B} + (\mathbf{A} \cdot \nabla)\mathbf{B}$, which gives

$$\begin{aligned} & \frac{\partial}{\partial t} (m_s n_s \mathbf{U}_s) + \nabla \cdot (m_s n_s \mathbf{U}_s \mathbf{U}_s) \\ &= m_s n_s \frac{\partial \mathbf{U}_s}{\partial t} + m_s \mathbf{U}_s \left[\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{U}_s) \right] + m_s n_s (\mathbf{U}_s \cdot \nabla) \mathbf{U}_s. \end{aligned} \quad (5.4.24)$$

From the continuity equation (5.4.6), the term in the brackets on the right side of the above equation is seen to be zero. The first moment of the Boltzmann equation then becomes

$$m_s n_s \left[\frac{\partial \mathbf{U}_s}{\partial t} + (\mathbf{U}_s \cdot \nabla) \mathbf{U}_s \right] = n_s e_s [\mathbf{E} + \mathbf{U}_s \times \mathbf{B}] - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_s + \frac{\delta_c \mathbf{P}_s}{\delta t}. \quad (5.4.25)$$

The quantity in the brackets on the left contains an operator that occurs frequently in fluid mechanics and is called the *convective derivative*. The term $\partial \mathbf{U}_s / \partial t$ gives the rate of change

of the velocity due to explicit time variations, and the term $(\mathbf{U}_s \cdot \nabla) \mathbf{U}_s$ gives the rate of change of the velocity due to spatial variations. If we follow a given volume element moving at velocity \mathbf{U}_s , it is easy to show that the convective derivative is simply the total derivative, $d\mathbf{U}_s/dt$, evaluated along the trajectory of the fluid element, that is,

$$\frac{d\mathbf{U}_s}{dt} = \frac{\partial \mathbf{U}_s}{\partial t} + (\mathbf{U}_s \cdot \nabla) \mathbf{U}_s. \quad (5.4.26)$$

Equation (5.4.25) can then be written in the form

$$m_s n_s \frac{d\mathbf{U}_s}{dt} = n_s e_s [\mathbf{E} + \mathbf{U}_s \times \mathbf{B}] - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_s + \frac{\delta_c \mathbf{P}_s}{\delta t}. \quad (5.4.27)$$

This equation is called the *momentum equation* and is easily recognized as Newton's law for the s th species of the plasma. The term on the left is the rate of change of the momentum per unit volume. The first term on the right is the electromagnetic force per unit volume, the second term is the force per unit volume due to pressure gradients, and the third term, $\delta_c \mathbf{p}_s / \delta t$, is the collisional drag force per unit volume.

5.4.3 Second moment

The second moment is obtained by multiplying the Boltzmann equation by the tensor $\mathbf{v}\mathbf{v}$ and integrating over all of velocity space. This procedure leads to a symmetric matrix with nine elements only six of which are independent, since the off-diagonal elements of a symmetric matrix are equal to each other. Here we will restrict our discussion to the trace of this matrix, which is obtained by multiplying the Boltzmann equation by the kinetic energy $(1/2) m_s v^2$ and integrating over velocity space, which gives

$$\begin{aligned} & \frac{m_s}{2} \int_V v^2 \frac{\partial f_s}{\partial t} d^3 v + \frac{m_s}{2} \int_V v^2 (\mathbf{v} \cdot \nabla f_s) d^3 v \\ & + \frac{1}{2} \int_V v^2 (\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3 v = \frac{m_s}{2} \int_V v^2 \frac{\delta_c f_s}{\delta t} d^3 v. \end{aligned} \quad (5.4.28)$$

The first integral on the left of the above equation can be written

$$\frac{m_s}{2} \int_V v^2 \frac{\partial f_s}{\partial t} d^3 v = \frac{\partial}{\partial t} \left(\int_V \frac{1}{2} m_s v^2 f_s d^3 v \right) = \frac{\partial W_s}{\partial t} \quad (5.4.29)$$

where W_s is the kinetic energy density. The second integral can be written

$$\frac{m_s}{2} \int_V v^2 (\mathbf{v} \cdot \nabla f_s) d^3 v = \nabla \cdot \left(\int_V \frac{1}{2} m_s \mathbf{v} v^2 f_s d^3 v \right) = \nabla \cdot \mathbf{Q}_s, \quad (5.4.30)$$

where

$$\mathbf{Q}_s = \int_V \frac{1}{2} m_s v^2 \mathbf{v} f_s d^3 v \quad (5.4.31)$$

is the kinetic energy flux. The third integral can be simplified by integrating by parts which, after discarding the contribution at infinite velocity, gives

$$\frac{1}{2} \int_V v^2 (\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3 v = -\frac{1}{2} \int_V f_s (\mathbf{F} \cdot \nabla_{\mathbf{v}} v^2) d^3 v. \quad (5.4.32)$$

Introducing the Lorentz force and noting that $\nabla_{\mathbf{v}} v^2 = 2\mathbf{v}$, and that $\mathbf{v} \cdot (\mathbf{v} \times \mathbf{B}) = 0$, the integral on the right side of the above equation can be written

$$\frac{1}{2} \int_V v^2 (\mathbf{F} \cdot \nabla_{\mathbf{v}} f_s) d^3 v = -e_s \mathbf{E} \cdot \int_V \mathbf{v} f_s d^3 v = -\mathbf{E} \cdot (n_s e_s \mathbf{U}_s) = -\mathbf{E} \cdot \mathbf{J}_s, \quad (5.4.33)$$

where \mathbf{J}_s is the current contributed by the species s . Using equations (5.4.29), (5.4.30), and (5.4.33) the second moment of the Boltzmann equation becomes

$$\frac{\partial W_s}{\partial t} + \nabla \cdot \mathbf{Q}_s - \mathbf{E} \cdot \mathbf{J}_s = \int_V \frac{1}{2} m_s v^2 \frac{\delta_c f_s}{\delta t} d^3 v. \quad (5.4.34)$$

This equation is called the *energy equation* because it is simply a statement of conservation of energy for the s th species. The term $\mathbf{E} \cdot \mathbf{J}_s$ is the contribution to the Joule heating by the s th species and the collision term on the right gives the rate of transfer of energy to the s th species due to collisions with other species. Summing the above equation over all species then gives an energy conservation equation for the entire plasma

$$\frac{\partial W}{\partial t} + \nabla \cdot \mathbf{Q} = \mathbf{E} \cdot \mathbf{J}, \quad (5.4.35)$$

where $W = \sum_s W_s$ is the total kinetic energy density, $\mathbf{Q} = \sum_s \mathbf{Q}_s$ is the total kinetic energy flux, and $\mathbf{E} \cdot \mathbf{J}$ is the Joule heating. In a fully ionized plasma, the collision terms sum to zero because for Coulomb collisions the effect of collisions is simply to transfer energy from one species to another without changing the overall energy.

5.4.4 The closure problem

It is evident by inspection of the first three moment equations that the n th moment equation always involves the $(n + 1)$ st moment of the distribution function. For example, the zeroth moment of the Boltzmann equation involves the first moment \mathbf{U}_s , etc. This hierarchy occurs because the $\mathbf{v} \cdot \nabla f_s$ term in the Boltzmann equation always increases the order of the moments by one. Therefore, for any finite number of moment equations, the number of unknown moments always exceeds the number of equations. There are never enough equations to solve for all of the unknown moments. This is called the closure problem. To obtain a closed set of equations it is then necessary to specify the $(n + 1)$ st moment in terms of the first n moments. For example, if only the continuity and momentum equations are used, a closed system of equations can be obtained by specifying an equation of state that relates the pressure tensor, $\overset{\leftrightarrow}{\mathbf{P}}_s$, to the number density, n_s . Unfortunately, this must be done in a purely ad hoc way, since there is no formal method to determine how the $(n +$

1)st moment should be related to the first n moments. Typically an equation of state is assumed that is based on one of the well-known equations of state used in thermodynamics. This approach works well when there are sufficient collisions to maintain the plasma in a state of local thermodynamic equilibrium, but often fails in a collisionless plasma. Next, we discuss some of the commonly used equations of state.

5.4.4.1 The cold plasma equation of state

In this simple approximation, the temperature is assumed to be so low that the pressure is negligible. The equation of state is simply $\overset{\leftrightarrow}{\mathbf{P}}_s = 0$ for any n_s . The momentum equation then becomes

$$m_s n_s \frac{d\mathbf{U}_s}{dt} = n_s e_s [\mathbf{E} + \mathbf{U}_s \times \mathbf{B}] + \frac{\delta_c \mathbf{P}_s}{\delta t}. \quad (5.4.36)$$

If the collisional drag term is assumed to be zero, the number density, n_s , can be factored out of this equation. The momentum equation then reduces to the equation of motion for a single particle,

$$m_s \frac{d\mathbf{U}_s}{dt} = e_s [\mathbf{E} + \mathbf{U}_s \times \mathbf{B}], \quad (5.4.37)$$

which is identical to the equation of motion (??) used in the analysis of waves in a cold plasma (Chapter ??). However, there is one subtle difference. In equation (5.4.37), \mathbf{U}_s is a fluid velocity field which is a function of the independent coordinates \mathbf{r} and t , whereas \mathbf{v}_s in equation (??) is the velocity of a single particle which is tagged and followed as a function of t . In some cases, this distinction can be important because the convective derivative, $d\mathbf{U}_s/dt$, consists of two terms, i.e.,

$$\frac{d\mathbf{U}_s}{dt} = \frac{\partial \mathbf{U}_s}{\partial t} + (\mathbf{U}_s \cdot \nabla) \mathbf{U}_s. \quad (5.4.38)$$

In the case of small amplitude waves in a fluid at rest (i.e., $\mathbf{U}_{s0} = 0$), the second term on the right can be ignored, since it involves the product of two first-order quantities, $(\mathbf{U}_{s1} \cdot \nabla) \mathbf{U}_{s1}$, which is small compared with the first-order quantity $\partial \mathbf{U}_{s1}/\partial t$. We then have $d\mathbf{U}_{s1}/dt \approx \partial \mathbf{U}_{s1}/\partial t$. The operator substitution $d/dt \rightarrow (-i\omega)$ is then a good first-order approximation. On the other hand, if the fluid is moving, i.e., $\mathbf{U}_{s0} \neq 0$, then the correct linearized form for the operator $d\mathbf{U}_{s1}/dt$ is

$$\frac{d\mathbf{U}_{s1}}{dt} = \frac{\partial \mathbf{U}_{s1}}{\partial t} + (\mathbf{U}_{s0} \cdot \nabla) \mathbf{U}_{s1}. \quad (5.4.39)$$

After Fourier transforming, this equation becomes

$$(-i\omega') \tilde{\mathbf{U}}_{s1} = (-i\omega + i\mathbf{k} \cdot \mathbf{U}_{s0}) \tilde{\mathbf{U}}_{s1}, \quad (5.4.40)$$

where ω' is the frequency in a frame of reference moving with the particle. From the above equation it follows that the frequency in the particle frame of reference is related to the

rest frame frequency by the equation $\omega' = \omega - \mathbf{k} \cdot \mathbf{U}_{s0}$. The term $\mathbf{k} \cdot \mathbf{U}_{s0}$ is the Doppler shift introduced by the change in the frame of reference. In general, the correct operator substitution for d/dt is $(-i\omega + i\mathbf{k} \cdot \mathbf{U}_{s0})$, and not $(-i\omega)$.

5.4.4.2 The adiabatic equation of state

The concept of an adiabatic equation of state is borrowed from thermodynamics. Using the ideal gas law, $PV = NkT$, and the first law of thermodynamics, $dQ = dU + P dV$, it can be shown that if no heat flows, then the pressure, P , and volume, V , are given by $PV^\gamma = \text{constant}$, where γ is the ratio of the heat capacity at constant pressure to the heat capacity at constant volume. The adiabatic equation of state applies to situations where the gas is compressed so rapidly that there is not enough time for heat to flow. The compression of a gas by a sound wave or by a rapidly moving piston are typical examples where the adiabatic equation of state can be used. Although the compression must occur sufficiently rapidly that the heat flow can be neglected, it must not occur so rapidly that collisions cannot maintain the gas in a state of local thermodynamic equilibrium. For example, the adiabatic equation of state cannot be used to describe the free expansion of a gas into an evacuated container.

Since local thermodynamic equilibrium is maintained by collisions, the question arises as to whether an adiabatic equation of state can be applied to a collisionless plasma. The obvious answer is no. However, microscopic wave-particle interactions sometimes play a role similar to collisions in an ordinary gas. The adiabatic equation of state may then provide an adequate approximation, even in a collisionless plasma.

When the adiabatic equation of state is used, it is usually assumed that collisions (or wave-particle interactions) maintain an isotropic velocity distribution. Under these conditions, the pressure tensor becomes diagonal, with equal diagonal elements, i.e.,

$$\overleftrightarrow{\mathbf{P}}_s = \overleftrightarrow{\mathbf{1}} P_s, \quad (5.4.41)$$

where $\overleftrightarrow{\mathbf{1}}$ is the unit tensor. Since we normally use the number density, n_s , rather than the volume, V , as the independent variable, it is easy to see that the adiabatic equation of state can be written in the following alternative form

$$\frac{d}{dt}(P_s/n_s^\gamma) = 0, \quad (5.4.42)$$

where the pressure obeys the ideal gas law, $P_s = n_s \kappa T_s$. From statistical mechanics it is well known that γ is determined by the number of degrees of freedom, f , and is given by

$$\gamma = \frac{f+2}{f}. \quad (5.4.43)$$

Thus, for a gas with three degrees of freedom ($f = 3$) it follows that $\gamma = 5/3$.

If the collision rate is too low to maintain an isotropic distribution then the number of degrees of freedom is controlled by the geometry of the compression process. Later, we

will show that in a collisionless plasma certain types of longitudinal waves behave as if the compression is one-dimensional, in which case $\gamma = 3$.

The power law index γ is also sometimes considered to be an adjustable parameter that can be selected to represent non-adiabatic equations of state. For example, $\gamma = 1$ represents an isothermal equation of state, and $\gamma = \infty$ represents an incompressible equation of state. In the fluid dynamics and astrophysics literature, the parameter γ is sometimes called the polytrope index [Fetter and Walecka, 1980].

5.4.4.3 The Chew–Goldberger–Low equation of state

If the collision rate in a magnetized plasma is too small to transfer momentum effectively between the parallel and perpendicular directions, an isotropic velocity distribution cannot be maintained. To account for such anisotropic effects, Chew, Goldberger, and Low [1956], often abbreviated CGL, developed a modified adiabatic equation of state in which the pressures parallel and perpendicular to the magnetic field are not the same. They showed that in a coordinate system in which the magnetic field is parallel to the z axis, the pressure tensor is diagonal and can be written in the form

$$\overset{\leftrightarrow}{\mathbf{P}}_s = \begin{bmatrix} P_{s\perp} & 0 & 0 \\ 0 & P_{s\perp} & 0 \\ 0 & 0 & P_{s\parallel} \end{bmatrix} \quad (5.4.44)$$

where the perpendicular and parallel pressures, $P_{s\perp}$ and $P_{s\parallel}$, are related to the magnetic field and number density by the equations

$$\frac{d}{dt} \left(\frac{P_{s\perp}}{n_s B} \right) = 0 \quad \text{and} \quad \frac{d}{dt} \left(\frac{P_{s\parallel} B^2}{n_s^3} \right) = 0. \quad (5.4.45)$$

The bi-Maxwellian velocity distribution given by equation (5.1.16) has a pressure tensor of the type described by equation (5.4.44).

The CGL equations of state (5.4.45) can be derived quite simply by making use of the 1st and 2nd adiabatic invariants discussed in Chapter ???. In a coordinate system at rest with respect to the plasma, the perpendicular pressure is given by the product of the number density and the average perpendicular kinetic energy, i.e., $P_{s\perp} = n_s \langle w_{s\perp} \rangle$. It then follows from the 1st adiabatic invariant and the relation $w_{s\perp} = \mu_s B$, that $P_{s\perp}/n_s B = \langle \mu_s \rangle$ is constant, which corresponds to the first equation in (5.4.45). The second equation can be derived similarly by making use of the 2nd adiabatic invariant. The parallel pressure is given by the product of the number density and the average parallel kinetic energy, i.e., $P_{s\parallel} = n_s \langle w_{s\parallel} \rangle$. Writing the 2nd adiabatic invariant in the form $v_{\parallel} L = \text{constant}$, where L is the distance between mirror points, it follows that $P_{s\parallel}$ is proportional to n_s / L^2 . The distance L between mirror points can next be eliminated by considering a volume with cross-sectional area A , perpendicular to the magnetic field, and length L , along the magnetic field. Conservation of mass implies $n_s A L = \text{constant}$, and conservation of magnetic flux implies $B A = \text{constant}$. Making use of these simple algebraic relations gives $P_{s\parallel} B^2 / n_s^3 = \text{constant}$, which is

the second equation in (5.4.45). Since the adiabatic invariants are constant only if the magnetic field varies slowly on a time scale compared to the parallel bounce period, the CGL equations of state can only be expected to be valid for time scales that are long compared to the bounce period.

5.5 Electron and ion pressure waves

To illustrate the use of the adiabatic equation of state, we next consider the propagation of small amplitude waves in a plasma with a finite pressure. Since the temporal fluctuations associated with a wave are relatively rapid, it is reasonable to assume that an adiabatic equation of state is valid. If we restrict our analysis to a plasma without a magnetic field, then the adiabatic equation of state can be written as a diagonal pressure tensor of the form (5.4.41), i.e.,

$$\overset{\leftrightarrow}{\mathbf{P}}_s = P_{s0} \left(\frac{n_s}{n_{s0}} \right)^\gamma \overset{\leftrightarrow}{\mathbf{1}}, \quad (5.5.1)$$

where P_{s0} and n_{s0} are the undisturbed pressure and density. For a small amplitude wave we assume, as in Chapter ??, that these quantities consist of a constant spatially uniform zero-order term plus a small first-order perturbation. The relevant equations are then

$$P_s = P_{s0} + P_{s1}, \quad (5.5.2)$$

$$n_s = n_{s0} + n_{s1}, \quad (5.5.3)$$

$$\mathbf{U}_s = \mathbf{U}_{s1}, \quad (5.5.4)$$

and

$$\mathbf{E} = \mathbf{E}_1, \quad (5.5.5)$$

where the zero-order velocity and electric field are assumed to be zero.

The equations that must be solved to determine the fluid motions are the continuity equation, the momentum equation, and the adiabatic equation of state:

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{U}_s) = 0, \quad (5.5.6)$$

$$m_s n_s \left[\frac{\partial \mathbf{U}_s}{\partial t} + (\mathbf{U}_s \cdot \nabla) \mathbf{U}_s \right] = n_s e_s \mathbf{E} - \nabla P_s, \quad (5.5.7)$$

and

$$P_s = P_{s0} \left(\frac{n_s}{n_{s0}} \right)^\gamma. \quad (5.5.8)$$

Linearizing and Fourier transforming equations (5.5.6)-(5.5.8), one obtains

$$(-i\omega) \tilde{n}_s = n_{s0} i \mathbf{k} \cdot \tilde{\mathbf{U}}_s = 0, \quad (5.5.9)$$

$$m_s n_{s0} (-i\omega) \tilde{\mathbf{U}}_s = n_{s0} e_s \tilde{\mathbf{E}} - i\mathbf{k} \mathbf{P}_s, \quad (5.5.10)$$

and

$$\tilde{P}_s = m_s \gamma_s C_s^2 \tilde{n}_s, \quad (5.5.11)$$

where for notational convenience we have dropped the subscript one on the first-order terms and introduced the thermal speed, C_s , of the s th species via the definition $C_s^2 = P_{s0}/m_s n_{s0}$. Note that since $P_{s0} = n_{s0} k T_{s0}$ it follows that $C_s^2 = \kappa T_{s0}/m_s$, which is consistent with the thermal speed defined by equation (??). Using equations (5.5.9) and (5.5.11), we can eliminate the first-order pressure from equation (5.5.10) to obtain

$$m_s (-i\omega) \tilde{\mathbf{U}}_s + m_s \gamma_s C_s^2 \frac{i\mathbf{k}}{\omega} (\mathbf{k} \cdot \tilde{\mathbf{U}}_s) = e_s \tilde{\mathbf{E}}. \quad (5.5.12)$$

It is useful to rewrite the above equation in matrix form,

$$\left(\overset{\leftrightarrow}{\mathbf{1}} - \gamma_s C_s^2 \frac{\mathbf{k}\mathbf{k}}{\omega^2} \right) \cdot \tilde{\mathbf{U}}_s = \frac{e_s}{(-i\omega)m_s} \tilde{\mathbf{E}}. \quad (5.5.13)$$

Since there is no preferred direction in this isotropic plasma, without loss of generality we can choose the \mathbf{k} vector to be aligned with the z axis. The above equation can then be written

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 - \gamma_s C_s^2 \frac{k^2}{\omega^2} \end{bmatrix} \begin{bmatrix} \tilde{U}_{sx} \\ \tilde{U}_{sy} \\ \tilde{U}_{sz} \end{bmatrix} = \frac{e_s}{(-i\omega)m_s} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}, \quad (5.5.14)$$

which can be solved for $\tilde{\mathbf{U}}_s$ to give

$$\begin{bmatrix} \tilde{U}_{sx} \\ \tilde{U}_{sy} \\ \tilde{U}_{sz} \end{bmatrix} = \frac{e_s}{(-i\omega)m_s} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1 - \gamma_s C_s^2 \frac{k^2}{\omega^2}} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}. \quad (5.5.15)$$

The conductivity tensor, defined by equation (??), can then be obtained by computing the current density $\tilde{\mathbf{J}} = \sum_s n_{s0} e_s \tilde{\mathbf{U}}_s$. Once the conductivity tensor has been determined, it is relatively simple to compute the dielectric tensor (??) and write down the homogeneous equation (??) for the electric field, which in matrix form is given by

$$\begin{bmatrix} -c^2 k^2 + \omega^2 - \omega_p^2 & 0 & 0 \\ 0 & -c^2 k^2 + \omega^2 - \omega_p^2 & 0 \\ 0 & 0 & \omega^2 - \sum_s \frac{\omega_{ps}^2}{1 - \gamma_s C_s^2 \frac{k^2}{\omega^2}} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (5.5.16)$$

The dispersion relation is then obtained by computing the determinant of the matrix, which is

$$D(k, \omega) = (-c^2 k^2 + \omega^2 - \omega_p^2)^2 \left(\omega^2 - \sum_s \frac{\omega_{ps}^2}{1 - \gamma_s C_s^2 \frac{k^2}{\omega^2}} \right) = 0. \quad (5.5.17)$$

It is evident from the above equation that there are two types of roots. The roots associated with the first term in parentheses correspond to transverse electromagnetic waves. The roots associated with the term in the large brackets correspond to longitudinal electrostatic

waves. It is interesting to note that the transverse electromagnetic mode is not affected by the pressure. This occurs because the electric field, $\tilde{\mathbf{E}} = (\tilde{E}_x, \tilde{E}_y, 0)$, and particle velocity, $\tilde{\mathbf{U}}_s = (\tilde{U}_{sx}, \tilde{U}_{sy}, 0)$, for the transverse electromagnetic mode are perpendicular to \mathbf{k} . From the continuity equation (5.5.9) it follows that the first-order number density, \tilde{n}_s , is zero. Since there is no density perturbation, there is no pressure perturbation associated with this mode. On the other hand, the longitudinal electrostatic mode is strongly affected by the pressure. For the longitudinal mode $\tilde{\mathbf{U}}_s$ and \mathbf{k} are parallel, so there is a non-zero fluctuation in the number density as well as the pressure. By factoring out ω^2 , the dispersion relation for the longitudinal mode can be written in the form

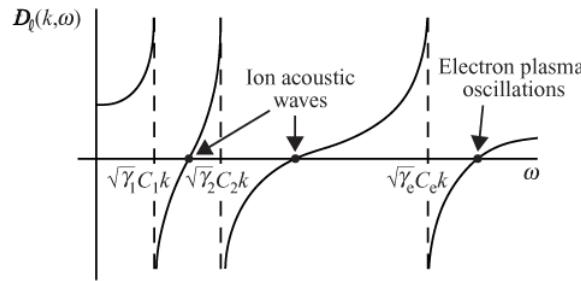


Fig. 5.4 A plot of $D_l(k, \omega)$ as a function of frequency for the longitudinal mode.

$$D_l(k, \omega) = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2 - \gamma_s C_s^2 k^2} = 0. \quad (5.5.18)$$

A plot of the function $D_l(k, \omega)$ is shown in Figure 5.4.

From this plot, it can be seen that the dispersion equation has one root associated with each species in the plasma. The root associated with the electrons is a modification of the electron plasma oscillation found in cold plasma theory and is often called the Langmuir mode, after Tonks and Langmuir [1929] who first studied this mode of propagation. As shown below, the electron pressure produces a shift in the frequency of the electron plasma oscillations. The roots associated with the ions have no analog in cold plasma theory and are called the ion acoustic modes. The properties of each of these modes are discussed below.

5.5.1 The Langmuir mode

Since the electron plasma frequency is much greater than the characteristic frequencies of the ions, for the Langmuir mode we can essentially ignore the motion of the ions. This is achieved by omitting the ion terms in the dispersion relation.

The dispersion relation then becomes

$$1 - \frac{\omega_{pe}^2}{\omega^2 - \gamma_e C_e^2 k^2} = 0, \quad (5.5.19)$$

or

$$\omega^2 = \omega_{pe}^2 + \gamma_e C_e^2 k^2. \quad (5.5.20)$$

This equation is called the Bohm-Gross dispersion relation, after Bohm and Gross [1949] who first derived it. If the electron temperature is zero, so that $C_e = 0$, the oscillation frequency reduces to the cold plasma result, $\omega^2 = \omega_{pe}^2$. It is easy to verify that the term $\gamma_e C_e^2 k^2$ comes from the electron pressure gradient term in the momentum equation (5.5.7). For long wavelengths ($k \approx 0$) the pressure gradient force is small compared with the electric field force. The oscillation frequency is then identical to the cold plasma result, i.e., $\omega^2 = \omega_{pe}^2$. As the wavelength decreases (and k increases), the pressure gradient force increases and eventually becomes larger than the electric field force. The pressure gradient force then causes the oscillation frequency to shift upwards relative to the cold plasma result.

Using the relation $\omega_{pe} \lambda_{De} = C_e$, we can rewrite the dispersion relation (5.5.20) in the form

$$\omega^2 = \omega_{pe}^2 \left[1 + \gamma_e \lambda_{De}^2 k^2 \right]. \quad (5.5.21)$$

The Debye length is now seen to be a basic length scale in the dispersion relation. The variation of the oscillation frequency with wave number is shown in Figure 5.5. As the wavelength approaches the Debye length, i.e., $k \approx 1/\lambda_{De}$, the oscillation frequency asymptotically approaches the dashed lines, which correspond to a phase velocity given by

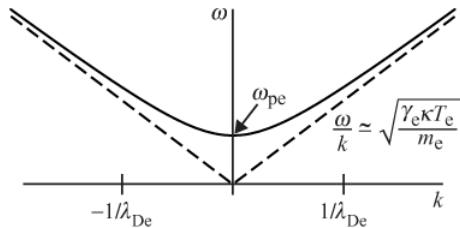


Fig. 5.5 The dispersion relation for electron plasma oscillations, including the effects of electron pressure.

$$\frac{\omega}{k} = \pm \sqrt{\frac{\gamma_e \kappa T_e}{m_e}}. \quad (5.5.22)$$

This equation is formally identical to the phase velocity of an acoustic wave in an ordinary gas, except the temperature is the electron temperature and the mass is the electron mass. Because of this close similarity, Langmuir waves are sometimes referred to as electron acoustic waves. However, this comparison is somewhat misleading. Later (in Chapter ??) we show that Langmuir waves are very strongly damped when $k\lambda_{De} \gtrsim 1$, due to a collisionless damping process called Landau damping. The failure of the moment equations to reveal this damping at large wave numbers is one of the shortcomings of the moment equations. The moment equations give a good description for $k\lambda_{De} \ll 1$, but fail when $k\lambda_{De} \gtrsim 1$.

5.5.2 The ion acoustic mode

To investigate the roots of the dispersion relation associated with the ions, we start by making the simplifying assumption that the ion temperature is small (i.e., $C_i = 0$). The dispersion relation (5.5.18) can then be written

$$D_\ell(k, \omega) = 1 - \frac{\omega_{\text{pi}}^2}{\omega^2} - \frac{\omega_{\text{pe}}^2}{\omega^2 - \gamma_e C_e^2 k^2} = 0. \quad (5.5.23)$$

For the ion roots, we expect that the phase velocity will be much less than the electron thermal velocity, so we assume that $\omega^2 \ll \gamma_e C_e^2 k^2$. The dispersion relation (5.5.23) can then be solved for ω^2 to yield

$$\omega^2 = \frac{1}{1 + \gamma_e \lambda_{\text{De}}^2 k^2} \left(\frac{\gamma_e \kappa T_e}{m_i} \right) k^2, \quad (5.5.24)$$

where we have again introduced λ_{De} by means of the relation $\omega_{\text{pe}} \lambda_{\text{De}} = C_e$, and substituted m_e/m_i for $\omega_{\text{pi}}^2/\omega_{\text{pe}}^2$. This mode of propagation is called the ion acoustic mode. Figure 5.6 shows a plot of the frequency ω of the ion acoustic mode as a function of the wave number, k . For long wavelengths $k \ll 1/\lambda_{\text{De}}$, ω varies linearly with k . In this region, $k \lambda_{\text{De}} \ll 1$, the phase velocity of the wave is essentially independent of wavelength and is given by

$$\frac{\omega}{k} = \pm \sqrt{\frac{\gamma_e \kappa T_e}{m_i}}. \quad (5.5.25)$$

Note that the phase velocity involves the *electron* temperature and the *ion* mass. This combination of parameters is interpreted as indicating that the electrons provide the pressure while the ions provide the inertia. As the wave number increases, the frequency begins to decrease relative to the linear relation given by equation (5.5.25) and asymptotically approaches the ion plasma frequency, ω_{pi} , for $k \lambda_{\text{De}} \gg 1$. It can be easily verified that the phase velocity of the ion acoustic mode is always much less than the electron thermal velocity if the ion temperature is zero. Therefore, the initial assumption that $\omega^2 \ll \gamma_e C_e^2 k^2$ is always satisfied.

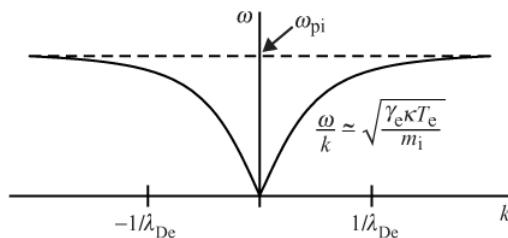


Fig. 5.6 The dispersion relation for the ion acoustic mode.

It is instructive to explore the electron and ion motions for the Langmuir mode and for the ion acoustic mode. From the momentum equation (5.5.10), it is easy to show that the

velocities of the electrons and ions are given, respectively, by

$$\tilde{U}_e = \left(\frac{-e}{-i\omega m_e} \right) \frac{\tilde{E}}{1 - \gamma_e C_e^2 \frac{k^2}{\omega^2}} \quad (5.5.26)$$

and

$$\tilde{U}_i = \left(\frac{+e}{-i\omega m_i} \right) \tilde{E}, \quad (5.5.27)$$

where we have again assumed that $C_i = 0$ in the ion equation. Taking the ratio of the electron-to-ion velocities, we obtain

$$\frac{\tilde{U}_e}{\tilde{U}_i} = - \left(\frac{m_i}{m_e} \right) \frac{1}{1 - \gamma_e C_e^2 \frac{k^2}{\omega^2}}. \quad (5.5.28)$$

For the Langmuir mode it is easy to verify that in the long wavelength limit the term $\gamma_e C_e^2 k^2 / \omega^2$ is much less than one. It follows then that $\tilde{U}_e / \tilde{U}_i \approx -m_i / m_e$. The minus sign indicates that the electrons and ions move in opposite directions. Also, since the ion mass is much greater than the electron mass, the amplitude of the electron motion is much greater than the amplitude of the ion motion. The ions are essentially motionless. In contrast, for the ion acoustic mode it is easy to show that in the long wavelength limit the term $\gamma_e C_e^2 k^2 / \omega^2$ is approximately m_i / m_e . It follows that $\tilde{U}_e / \tilde{U}_i \approx 1 + m_e / m_i$. The electrons and ions then move in the same direction at nearly the same velocity, causing a fluid-like motion that is very similar to an acoustic wave in an ordinary gas. The small difference in the electron and ion velocities caused by the m_e / m_i term produces the charge separation that is responsible for the electric field of the wave. It is this electric field that couples the electron and ion motions.

If the electron and ion temperatures are equal ($T_e = T_i$) it can be shown that the phase velocity of the ion acoustic mode is comparable to the ion thermal speed. The ion acoustic mode then suffers heavy Landau damping due to absorption of wave energy by the ions, very similar to the situation for electron plasma oscillations when $k\lambda_{De} \gtrsim 1$. Again, this damping is not predicted by the moment equations. If $T_e \gg T_i$ equation (5.5.24) shows that the phase velocity of the ion acoustic wave is much larger than the ion thermal speed. Later, in Chapter ??, we show that it is only when $T_e \gg T_i$ that the ion acoustic wave can propagate with relatively little damping. This justifies our earlier assumption that the ion temperature must be small (i.e., $C_i = 0$).

5.6 Collisional drag force

So far we have ignored the effect of collisions. The effect of collisions enters into the momentum equations (5.4.27) via the collisional drag force, $\delta_c \mathbf{p}_s / \delta_t$. For a collisionless plasma this term is often assumed to be zero. If collisions cannot be neglected then an equation is needed for the collisional drag force. Collisional effects are quite complicated and a full discussion is not presented until Chapter ???. However, for some applications a simple model called the Lorentz gas model can be used. This model is discussed below.

5.6.1 The Lorentz gas model

In the Lorentz gas model it is assumed that the plasma particles are scattered by fixed immobile scattering centers. If, in addition, we assume that the scattering is isotropic, the average momentum change per collision is simply $-mv$, since the average momentum after a collision is zero. The average drag force per unit volume for a particle of type s is then given by

$$\frac{\delta_c \mathbf{p}_s}{\delta t} = v_s \int_V (-m_s \mathbf{v}) f_s d^3 v, \quad (5.6.1)$$

or

$$\frac{\delta_c \mathbf{p}_s}{\delta t} = -v_s m_s n_s \mathbf{U}_s, \quad (5.6.2)$$

where v_s is an effective collision frequency. This equation shows that the collisional drag force $\delta_c \mathbf{p}_s / \delta t$, is directly proportional to the average momentum density, $m_s n_s \mathbf{U}_s$, with the collision frequency, v_s , acting as the constant of proportionality.

The Lorentz gas model is particularly well suited for electron scattering by neutral particles. Neutral particles essentially act as hard sphere scattering centers, thereby causing nearly isotropic scattering. Since the electron mass is much less than the mass of the neutral particles, very little momentum is transferred to the neutral particle, so the assumption of fixed immobile scattering centers is valid. For ion-neutral collisions the isotropic scattering approximation is not as good, since an appreciable amount of momentum is transferred to the neutral atom during the collision, thereby introducing a non-isotropic component to the scattering. Nevertheless, the Lorentz gas model is still sometimes used for ion-neutral collisions. Even though the scattering is not completely isotropic, it can be shown that the drag force is, to a first approximation, proportional to the relative velocity between the ions and the neutral atoms. For scattering by charged particles, the isotropic assumption is not valid since charged-particle scattering is highly anisotropic, see Section ???. For a more detailed discussion of collisions between charged particles, see Chapter ??.

5.6.2 Ohm's law

When a steady electric field is applied to a plasma, the electrons and ions are accelerated in opposite directions by the electric field force, thereby producing a current. In the absence of collisions the acceleration would continue indefinitely. It would then not be possible to define a conductivity, since a steady-state current is never achieved. However, if a collisional drag force is present, an equilibrium velocity is eventually reached at which the drag force balances the electric field force. If the equilibrium velocity is proportional to the electric field, then a conductivity, σ , can be defined by means of Ohm's law, $\mathbf{J} = \sigma \mathbf{E}$. To evaluate this conductivity, consider a homogeneous plasma with no magnetic field. The time required to reach equilibrium can be determined by assuming that a uniform electric field, \mathbf{E}_0 , is turned on at $t = 0$. Using equation (5.6.2) for the collisional drag force, the momentum equation then becomes

$$m_s n_s \frac{\partial \mathbf{U}_s}{\partial t} = e_s n_s \mathbf{E}_0 - v_s m_s n_s \mathbf{U}_s \quad (5.6.3)$$

which simplifies to

$$\frac{\partial \mathbf{U}_s}{\partial t} + v_s \mathbf{U}_s = \frac{e_s}{m_s} \mathbf{E}_0. \quad (5.6.4)$$

This linear differential equation is easily solved for the drift velocity \mathbf{U}_s . The solution is

$$\mathbf{U}_s \frac{e_s \mathbf{E}_0}{m_s v_s} (1 - e^{-v_s t}) \quad (5.6.5)$$

The drift velocity starts at zero and exponentially approaches a steady-state equilibrium on a time scale of $\tau = 1/v_s$. The initial phase of the transient response is controlled by inertia and the asymptotic final phase is controlled by the collisional drag force. For times $t \gg \tau$ it is easy to see from equation (5.6.5) that the steady-state drift velocity is given by $\mathbf{U}_s = e_s \mathbf{E}_0 / m_s v_s$. The resulting current is then given by

$$\mathbf{J}_s = n_s e_s \mathbf{U}_{s0} = \left(\frac{n_s e_s^2}{m_s v_s} \right) \mathbf{E}_0, \quad (5.6.6)$$

where the quantity in parentheses is the conductivity due to the s th species. The total conductivity is then the sum of the conductivities of each species

$$\sigma = \sum_s \sigma_s \quad (5.6.7)$$

where

$$\sigma_s = \frac{n_s e_s^2}{m_s v_s}. \quad (5.6.8)$$

Because the electron mass is much smaller than the ion mass, electrons usually provide the main contribution to the conductivity, so $\sigma \approx \sigma_e$.

An Ohm's law conductivity, σ , such as given by equation (5.6.8), arises whenever the drift velocity is linearly proportional to the electric field. Note that in contrast to the conductivity encountered in the analysis of waves in a cold plasma, which was always imaginary, the conductivity caused by collisions is real, which implies energy is being dissipated. The energy dissipation occurs because the ordered motion produced by the applied electric field is converted into random thermal motion (i.e., heat) by collisions. Because of their small mass, electrons are heated more rapidly than ions during such collisional interactions. For fusion plasmas this is unfortunate, since it is the ions that must be heated to a high temperature in order to initiate a fusion reaction. Therefore, collisional or "Ohmic" heating is a relatively ineffective method of heating a fusion plasma.

5.6.3 Pedersen and Hall conductivities

If a magnetic field is present, the conductivity is modified because of the $e_s \mathbf{U}_s \times \mathbf{B}$ force. For a homogeneous plasma with a magnetic field \mathbf{B}_0 and an electric field \mathbf{E}_0 , the momentum equation in steady state is given by

$$0 = n_s e_s [\mathbf{E}_0 + \mathbf{U}_s \times \mathbf{B}_0] - v_s m_s n_s \mathbf{U}_s. \quad (5.6.9)$$

Assuming that the magnetic field is parallel to the z axis, the above equation can be written in matrix form as follows:

$$\begin{bmatrix} 1 & -\frac{\omega_{cs}}{v_s} & 0 \\ \frac{\omega_{cs}}{v_s} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} U_{sx} \\ U_{sy} \\ U_{sz} \end{bmatrix} = \frac{e_s}{m_s v_s} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}, \quad (5.6.10)$$

where $\omega_{cs} = e_s B_0 / m_s$. This equation can be inverted to give

$$\begin{bmatrix} U_{sx} \\ U_{sy} \\ U_{sz} \end{bmatrix} = \frac{e_s}{m_s v_s \left(1 + \frac{\omega_{cs}^2}{v_s^2}\right)} \begin{bmatrix} 1 & \frac{\omega_{cs}}{v_s} & 0 \\ -\frac{\omega_{cs}}{v_s} & 1 & 0 \\ 0 & 0 & 1 + \frac{\omega_{cs}^2}{v_s^2} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}. \quad (5.6.11)$$

By computing the current using $\mathbf{J} = \sum_s n_s e_s \mathbf{U}_s$ and expressing the current in the form $\mathbf{J} = \vec{\sigma} \cdot \mathbf{E}$, the conductivity tensor can be easily identified and is given by

$$\vec{\sigma} \begin{bmatrix} \sigma_{\perp} & \sigma_H & 0 \\ -\sigma_H & \sigma_{\perp} & 0 \\ 0 & 0 & \sigma_{\parallel} \end{bmatrix}, \quad (5.6.12)$$

where

$$\sigma_{\perp} = \sum_s \frac{\sigma_s}{\left(1 + \omega_{cs}^2/v_s^2\right)}, \quad \sigma_H = \sum_s \frac{\sigma_s (\omega_{cs}/v_s)}{\left(1 + \omega_{cs}^2/v_s^2\right)} \quad (5.6.13)$$

and $\sigma_{\parallel} = \sum_s \sigma_s$.

The conductivity tensor elements σ_{\parallel} and σ_{\perp} are called the parallel and perpendicular (Pedersen) conductivities, respectively, and the element σ_H is called the Hall conductivity. For currents flowing parallel to the magnetic field, the magnetic field has no effect on the conductivity since the $e_s \mathbf{U}_s \times \mathbf{B}$ force is zero. Therefore, the parallel conductivity is the same as in a plasma with no magnetic field. The perpendicular conductivity gives the conductivity perpendicular to the magnetic field. The magnetic field has a strong effect on the perpendicular conductivity, particularly when $\omega_{cs} \gg v_s$. The Hall conductivity is the off-diagonal element in the conductivity tensor. For an electric field applied perpendicular to the magnetic field, the Hall conductivity causes a current to flow perpendicular to both \mathbf{E}_{\perp} and \mathbf{B}_0 . This component of the current is called the Hall current. The Hall current is caused by the $e_s \mathbf{U}_s \times \mathbf{B}$ force in the momentum equation, which is in the opposite direction for positive and negative charges, thereby causing a current to flow perpendicular to the applied electric field. Because of the Hall current, the total current flows at an angle to the electric field, as shown in Figure 5.7. It is relatively straightforward to show that the angle between the perpendicular electric field \mathbf{E}_{\perp} and the perpendicular component of the current, $\mathbf{J}_{s\perp}$, is given by

$$\tan \theta = \frac{J_{sy}}{J_{sx}} = -\frac{\omega_{cs}}{v_s}. \quad (5.6.14)$$

From the above equation one can see that the ratio of the cyclotron frequency to the collision frequency plays an important role in determining the direction of current flow. If

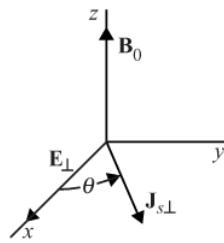


Fig. 5.7 In a magnetized plasma collisions cause the current to flow at an angle to the electric field.

$\omega_{cs} \ll v_s$ then the conductivity tensor becomes diagonal with $\hat{\sigma} = \hat{1}\sigma$, where σ is given by equation (5.6.7). On the other hand, if $\omega_{cs} \gg v_s$ then σ_\perp and σ_H go to zero, and σ_\parallel becomes infinite. The only current that can flow under these conditions is along the magnetic field.

5.7 Ambipolar diffusion

Frequently in laboratory and space applications one finds situations where plasma is produced by an ionization process in one region and transported to some other region where it is lost. If the transport is caused by random collisional forces, the transport process is called diffusion. Diffusion causes particles to move from a region of high density to a region of low density. In this section we consider the diffusion of electrons and one species of positive ions using the Lorentz gas model. This model is often applied to the diffusion of electrons and ions through a neutral gas.

If the diffusion is a steady-state process ($\partial/\partial t \approx 0$), the momentum equation for the electrons and ions can be written

$$m_s n_s (\mathbf{U}_s \cdot \nabla) \mathbf{U}_s = n_s e_s [\mathbf{E} + \mathbf{U}_s \times \mathbf{B}] - \nabla P_s - v_s m_s n_s \mathbf{U}_s. \quad (5.7.1)$$

Since diffusion is associated with a density gradient, the inhomogeneous terms cannot be neglected. Usually the drift velocity is much smaller than the root-mean-square thermal velocity (i.e., $U_s^2 \ll \kappa T_s/m_s$). Under these conditions the inertial term on the left-hand side of the above equation is small compared to the pressure gradient term. Dropping the inertial term and assuming that the magnetic field is zero, the momentum equation can be solved for the particle flux, $n_s \mathbf{U}_s$, which is given by

$$n_s \mathbf{U}_s = -\frac{1}{v_s m_s} \nabla P_s + \frac{e_s n_s}{v_s m_s} \mathbf{E}. \quad (5.7.2)$$

Substituting the ideal gas law, $P_s = n_s \kappa T_s$, for the pressure term, the above equation can be written

$$n_s \mathbf{U}_s = -\nabla(\bar{D}_s n_s) + \bar{\mu}_s n_s \mathbf{E}, \quad (5.7.3)$$

where the quantities $\bar{D}_s = \kappa T_s/v_s m_s$ and $\bar{\mu}_s = e_s/v_s m_s$ are called the diffusion and mobility coefficients. The diffusion coefficient provides a measure of the transport rate caused by the gradient in the number density, and the mobility coefficient provides a measure of the transport rate caused by the electric field.

To compute the particle flux, we must also take into account the constraint imposed by the continuity equation. If we avoid source or loss regions, there are two continuity equations, one for the electrons

$$\frac{\partial n_e}{\partial t} + \nabla \cdot (n_e \mathbf{U}_e) = 0, \quad (5.7.4)$$

and one for the ions

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{U}_i) = 0. \quad (5.7.5)$$

Subtracting the two continuity equations gives

$$\nabla \cdot [n(\mathbf{U}_e - \mathbf{U}_i)] = \frac{\partial}{\partial t}(n_e - n_i) \quad (5.7.6)$$

where $n = n_e \approx n_i$. If the plasma is quasi-neutral, so that $n_e \approx n_i$, it follows that the right-hand side of the above equation is very small, so that to a good approximation

$$\nabla \cdot [n(\mathbf{U}_e - \mathbf{U}_i)] \approx 0. \quad (5.7.7)$$

In many cases it can be shown that $\mathbf{U}_e = \mathbf{U}_i$ provides the only solution of the above equation. For example, for a cylindrically symmetric plasma where all the parameters are independent of the ϕ and z coordinates, equation (5.7.7) becomes

$$\frac{1}{\rho} \frac{d}{d\rho} [\rho n(U_{\rho e} - U_{\rho i})] = 0, \quad (5.7.8)$$

where ρ is the radial distance from the central axis. The above equation implies that the term $e\rho n(U_{\rho e} - U_{\rho i})$ must be constant. If there is no current source at $\rho = 0$ then it follows that $U_{\rho e} = U_{\rho i}$. These drift velocities depend on the axial current. If the boundary conditions at the ends of the cylinder are such that no current can flow along the z axis, then it follows that $\mathbf{U}_e = \mathbf{U}_i$. A process for which $\mathbf{U}_e = \mathbf{U}_i$ is called ambipolar diffusion.

If ambipolar diffusion is assumed, so that $\mathbf{U}_e = \mathbf{U}_i = \mathbf{U}$, then the following two equations must be satisfied

$$n\mathbf{U} = -\nabla(\bar{D}_e n) - |\bar{\mu}_e|n\mathbf{E}, \quad (5.7.9)$$

and

$$n\mathbf{U} = -\nabla(\bar{D}_i n) + \bar{\mu}_i n\mathbf{E}. \quad (5.7.10)$$

The negative sign in the electron mobility term has been used to emphasize the fact that the

electric field force on the electrons is directed opposite to the electric field. The electric field in these equations is intimately related to the ambipolar diffusion process. In the absence of an electric field, the diffusion rate of the electrons would be much higher than the diffusion rate of the ions because of the lighter mass and higher thermal velocity of the electrons. However, as soon as the electrons start to diffuse, a polarization charge develops. This polarization charge produces an electric field that reduces the electron flux and enhances the ion flux. Equilibrium is achieved when the electron and ion drift velocities are equal. Equations (5.7.9) and (5.7.10) are self-consistent relations that must be satisfied for this equilibrium to be achieved. The electric field can be eliminated from these equations to give

$$n\mathbf{U} = -\nabla(\bar{D}_a n), \quad (5.7.11)$$

where

$$\bar{D}_a = \frac{\bar{D}_e \bar{\mu}_i + \bar{D}_i |\bar{\mu}_e|}{\bar{\mu}_i + |\bar{\mu}_e|}. \quad (5.7.12)$$

The quantity \bar{D}_a is called the ambipolar diffusion coefficient. Equation (5.7.11) can then be substituted into the continuity equation to obtain

$$\frac{\partial n}{\partial t} = \nabla^2(\bar{D}_a n). \quad (5.7.13)$$

This equation is of a general form called the diffusion equation.

The magnitude of the ambipolar diffusion coefficient can be estimated by evaluating the various terms in \bar{D}_a . Because the electron mobility is always much greater than the ion mobility, $|\bar{\mu}_e| \gg \bar{\mu}_i$, it is easy to show that to a good approximation

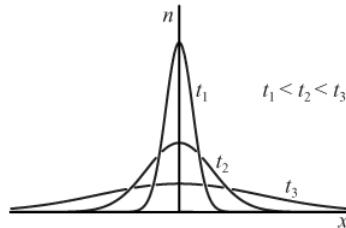


Fig. 5.8 Solutions to the diffusion equations at a sequence of times t_1, t_2 , and t_3 .

$$\bar{D}_a \approx \bar{D}_i \left(1 + \frac{T_e}{T_i} \right). \quad (5.7.14)$$

For equal electron and ion temperatures, the ambipolar diffusion coefficient is approximately twice the ion diffusion coefficient, which is much smaller than the electron diffusion coefficient. Thus, even though the electrons tend to diffuse much more rapidly than the ions, the ambipolar electric field restricts the electrons to diffuse at the same rate as ions. Therefore, the ion diffusion effectively controls the electron transport.

It is relatively easy to show that solutions of the diffusion equation (5.7.13) have the property of eliminating peaks in the density distribution. This is because the flux of particles is always away from regions of higher density. A time sequence of solutions with a constant diffusion coefficient is shown in Figure 5.8. From simple scaling considerations, it can be shown that the time scale for the decay is $\tau_D = L^2/\bar{D}_a$, where L is a typical length scale. In steady state ($\partial/\partial t = 0$), the diffusion equation reduces to Laplace's equation

$$\nabla^2(\bar{D}_a n) = 0, \quad (5.7.15)$$

the solutions of which are discussed in most textbooks on electricity and magnetism.

Problems

5.1. The general form of a Maxwellian velocity distribution drifting at velocity \mathbf{U}_0 is

$$f(\mathbf{v}) = n_0 \left(\frac{m}{2\pi\kappa T} \right)^{3/2} \exp \left[-\frac{m(\mathbf{v} - \mathbf{U}_0)^2}{2\kappa T} \right].$$

(a) Show that

$$\int_{-\infty}^{\infty} f(\mathbf{v}) d^3 v = n_0.$$

(b) Show that

$$\langle \mathbf{v} \rangle = \frac{1}{n} \int_{-\infty}^{\infty} \mathbf{v} f(\mathbf{v}) d^3 v = \mathbf{U}_0.$$

(c) Show that the pressure tensor is

$$\overset{\leftrightarrow}{\mathbf{P}} = \begin{vmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{vmatrix}, \quad \text{where } P = n_0 \kappa T.$$

5.2. Later we will use a one-dimensional distribution function of the form

$$F_0 = \frac{C}{\pi} \frac{1}{C^2 + v_z^2}$$

called the Cauchy distribution. Show that $\int_{-\infty}^{\infty} F_0 dv_z = 1$.

5.3. The distribution function used in problem 5.2 above has the disadvantage that the second moment is infinite, which implies an infinite temperature. To avoid this problem, one can use a function of the form $1/(C^2 + v_z^2)^2$, so that

$$F_0(v_z) = \frac{2C^3}{\pi} \frac{1}{(C^2 + v_z^2)^2}.$$

(a) Show that

$$\int_{-\infty}^{\infty} F_0(v_z) dv_z = 1.$$

(b) Show that the root-mean-square value of v_z is C . In other words that

$$\langle v_z^2 \rangle^{1/2} = C.$$

5.4. To represent a loss-cone distribution, we sometimes use a distribution function of the form

$$f_0(v_{\perp}, v_{\parallel}) = \frac{n_0}{(2\pi)^{3/2} C_{\perp}^2 C_{\parallel} \ell!} \left(\frac{v_{\perp}^2}{2C_{\perp}^2} \right)^{\ell} \exp\left(-\frac{v_{\perp}^2}{2C_{\perp}^2}\right) \exp\left(\frac{v_{\parallel}^2}{2C_{\parallel}^2}\right)$$

where $C_{\perp}^2 = \kappa T_{\perp}/m$ and $C_{\parallel}^2 = \kappa T_{\parallel}/m$, and ℓ is an integer.

(a) Show that

$$\int_{-\infty}^{\infty} f_0(\mathbf{v}) d^3v = n_0.$$

(b) Show that

$$\overset{\leftrightarrow}{\mathbf{P}} = \begin{vmatrix} P_{\perp} & 0 & 0 \\ 0 & P_{\perp} & 0 \\ 0 & 0 & P_{\parallel} \end{vmatrix},$$

where $P_{\perp} = (\ell + 1)n_0\kappa T_{\perp}$ and $P_{\parallel} = n_0\kappa T_{\parallel}$.

5.5. As an example of a solution of the Vlasov equation based on a constant of the motion, we discussed a Maxwellian

$$f(\mathbf{v}) = n_0 \left(\frac{m}{2\pi\kappa T_0} \right)^{3/2} \exp\left[\frac{-(\frac{1}{2}mv^2 + q\Phi)}{\kappa T_0} \right]$$

where Φ is the electrostatic potential, $\mathbf{E} = -\nabla\Phi$. Explicitly show that this distribution function satisfies the Vlasov equation. In other words, show that

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \frac{q\mathbf{E}}{m} \cdot \nabla_{\mathbf{v}} f = 0.$$

5.6. As an example of the use of the above solution, consider a plasma of electrons and ions that have distribution functions

$$f_e = n_0 \left(\frac{m_e}{2\pi\kappa T_e} \right)^{3/2} \exp\left[\frac{-(\frac{1}{2}m_e v^2 - e\Phi)}{\kappa T_e} \right],$$

and

$$f_i = n_0 \left(\frac{m_i}{2\pi\kappa T_i} \right)^{3/2} \exp \left[\frac{-(\frac{1}{2}m_i v^2 + e\Phi)}{\kappa T_i} \right],$$

- (a) By integrating over velocity space to obtain the number densities n_e and n_i , show that the potential must satisfy

$$\nabla^2 \Phi = -\frac{n_0 e}{\epsilon_0} \left[\exp \left(\frac{-e\Phi}{\kappa T_i} \right) - \exp \left(\frac{e\Phi}{\kappa T_e} \right) \right].$$

- (b) By assuming that $e\Phi/\kappa T_i \ll 1$ and $e\Phi/\kappa T_e \ll 1$ show that

$$\nabla^2 \Phi = \frac{n_0 e^2}{\epsilon_0 \kappa} \left[\frac{1}{T_i} + \frac{1}{T_e} \right] \Phi.$$

- (c) A charge Q is introduced at $r = 0$. Show that the potential is given by

$$\Phi = \frac{1}{4\pi\epsilon_0} \frac{Q}{r} e^{-r/\lambda_D},$$

where $(1/\lambda_D)^2 = (1/\lambda_{De})^2 + (1/\lambda_{Di})^2$.

- (d) In the process of integrating f_e and f_i over all of velocity space, we have made an error. Identify and discuss the nature of this error. Two cases must be considered, $Q > 0$ and $Q < 0$.

- 5.7. A plasma with a magnetic field $\mathbf{B}_0(s)$ has a parallel electric field that can be described by a monotonic electrostatic potential $\Phi(s)$, where s is the distance along the magnetic field line. Suppose that an absorbing boundary (the wall of the plasma chamber or the atmosphere) exists at $s = m$.

- (a) By assuming conservation of energy and conservation of the first adiabatic invariant show that, at an arbitrary point s , the loss cone in velocity space $(v_{||}, v_{\perp})$ is given by

$$v_{||}^2 - \left(\frac{B_m}{B} - 1 \right) v_{\perp}^2 = \frac{2q}{m} (\Phi_m - \Phi).$$

- (b) Sketch the shape of the loss cone for electrons and ions assuming that $\Phi_m > \Phi$ and $B_m > B$.

- (c) Suppose ions are now emitted at $s = m$. What region of $(v_{||}, v_{\perp})$ is accessible to these ions again assuming that $\Phi_m > \Phi$ and $B_m > B$.

- (d) Find the ion distribution function at the arbitrary point if the distribution function at $s = m$ is a Maxwellian

$$f_i = n_0 \left(\frac{m}{2\pi\kappa T} \right)^{3/2} \exp \left[-\frac{mv^2}{2\kappa T} \right].$$

- 5.8. Show that the Vlasov equation in cylindrical phase-space coordinates $(\rho, \phi, z, \dot{\rho}, \dot{\phi}, \dot{z})$ is given by

$$\frac{\partial f}{\partial t} + \dot{\rho} \frac{\partial f}{\partial \rho} + \phi \frac{\partial f}{\partial \phi} + z \frac{\partial f}{\partial z} + \rho \phi^2 \frac{\partial f}{\partial \dot{\rho}} - \frac{2\dot{\rho}\phi}{\rho} \frac{\partial f}{\partial \dot{\phi}} + \frac{1}{m} \left(F_\rho \frac{\partial f}{\partial \rho} + \frac{F_\phi}{\rho} \frac{\partial f}{\partial \phi} + F_z \frac{\partial f}{\partial z} \right) = 0.$$

5.9. Show that the following distribution function

$$f\left(\frac{1}{2}mv^2, m\rho^2\dot{\phi} + q\rho A_\phi\right)$$

is a time stationary solution of the Vlasov equation for an azimuthally symmetric magnetic field of the form

$$\mathbf{B} = \frac{1}{\rho} \frac{\partial}{\partial \rho} [\rho A_\phi(\rho)] \hat{\mathbf{z}}.$$

References

- Bohm, D., and Gross, E. P. 1949. Theory of plasma oscillations. A. Origin of medium-like behavior; B. Excitation and damping of oscillations. *Phys. Rev.* **75**, 1851–1876.
- Boltzmann, L. 1896–1898. *Vorlesungen Über Gastheorie, 1896-1898*. Translated by S. Brush, *Lectures on Gas Theory, 1896-1898*. University of California Press, reprinted by Dover Publications: New York, 1995.
- Chew, G. F., Goldberger, M. L., and Low, F. E. 1956. The Boltzmann equation and the one-fluid hydrodynamic equations in the absence of collisions. *Proc. R. Soc. London, Ser. A* **236**, 112–118.
- Fetter, A. L., and Walecka, J. D. 1980. *Theoretical Mechanics of Particles and Continua*. McGraw-Hill: New York, p. 345.
- Kaplan, W. 1952. *Advanced Calculus*. Addison-Wesley: Reading, MA, p. 99.
- Tonks, L., and Langmuir, I. 1929. Oscillations in ionized gases. *Phys. Rev.* **33**, 195–210.
- Vlasov, A. A. 1945. On the kinetic theory of an assembly of particles with collective interaction. *J. Phys. (USSR)* **9**, 25–40.

Further reading

- Chen, F. F. *Introduction to Plasma Physics and Controlled Fusion*, Vol. 1: *Plasma Physics*. Plenum Press: New York, 1984 (reprinted in 1990), Chapter 7.
- Montgomery, D. C., and Tidman, D. A. *Plasma Kinetic Theory*. McGraw-Hill: New York, 1964, Chapter 1.
- Nicholson, D. R. *Introduction to Plasma Theory*. Krieger Publishing: Malabar, FL, 1992 (originally published by Wiley: New York, 1983), Chapters 3,4, 5, and 6.
- Parks, G. K. *Physics of Space Plasmas: An Introduction*. Addison-Wesley: Redwood City, CA, 1991 (reprinted in 2000), Chapter 2.

6

Magnetohydrodynamics

Historically magnetohydrodynamics, abbreviated MHD, preceded the development of modern plasma physics. The original intent of MHD was to treat a plasma as a conducting fluid [see Cowling, 1957]. The governing equations were adapted from fluid mechanics with appropriate modifications to account for electrical forces. To obtain a complete set of equations, it was necessary to specify the current as a function of the applied electric field. This was accomplished by using a linear Ohm's law, such as is often used to describe conducting media. Since, to a first approximation, plasmas are electrically neutral, the net charge density was assumed to be identically zero. Also, since fluid motions tend to be slow compared to the characteristic time scales of a plasma, the displacement current was assumed to be small compared to the conduction current. These assumptions, together with an appropriate equation of state, were sufficient to obtain a closed system of equations.

Although it would be adequate simply to write down the equations of MHD as they were originally postulated, it is useful to try to derive the equations from first principles, using the moment equations developed in Chapter ???. Although this procedure will only be partially successful, it has the advantage of revealing more clearly the underlying assumptions and the range of applicability of MHD. It also has the advantage of providing a theoretical basis for precisely defining certain quantities, such as the fluid velocity and the plasma pressure.

6.1 The basic equations of MHD

To derive the MHD equations from the moment equations, it is convenient to define the mass density as the sum of the mass densities of the individual species

$$\rho_m = \sum_s \rho_{ms}, \quad (6.1.1)$$

where $\rho_{ms} = m_s n_s$, and the fluid velocity as the mass weighted average of the velocities of the individual species

$$\mathbf{U} = \frac{1}{\rho_m} \sum_s \rho_{ms} \mathbf{U}_s, \quad (6.1.2)$$

where \mathbf{U}_s is defined by equation (??). Note that since the ions have a much greater mass than the electrons, the fluid velocity \mathbf{U} is heavily weighted by the velocity of the ions.

6.1.1 The mass continuity equation

To derive the mass continuity equation for an MHD fluid, we start with the mass continuity equation for the s th species

$$\frac{\partial \rho_{ms}}{\partial t} + \nabla \cdot (\rho_{ms} \mathbf{U}_s) = 0, \quad (6.1.3)$$

which follows by multiplying equation (??) by m_s and using the definition $\rho_{ms} = m_s n_s$. Summing the above equation over all species s and making use of the definitions for the mass density (6.1.1) and the fluid velocity (6.1.2), we obtain the equation

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \mathbf{U}) = 0. \quad (6.1.4)$$

This equation is identical to the well-known mass continuity equation of fluid dynamics [see Batchelor, 1967]. Note that the above derivation depends critically on the fact that the fluid velocity is the mass weighted average of the flow velocities of the individual species. A simple average of the flow velocities of the individual species would not give this result.

6.1.2 The momentum equation

To derive an equation for the time rate of change of the fluid momentum, we start with the momentum equation (??) for the s th species. Summing this equation over all species, we obtain

$$\begin{aligned} & \sum_s \frac{\partial}{\partial t} (\rho_{ms} \mathbf{U}_s) + \sum_s \nabla \cdot (\rho_{ms} \mathbf{U}_s \mathbf{U}_s) \\ &= \rho_q \mathbf{E} + \mathbf{J} \times \mathbf{B} - \nabla \cdot \left(\sum_s \overset{\leftrightarrow}{\mathbf{P}}_s \right) + \sum_s \frac{\delta_c \mathbf{p}_s}{\delta t}, \end{aligned} \quad (6.1.5)$$

where we have used $\rho_q = \sum_s n_s e_s$ and $\mathbf{J} = \sum_s n_s e_s \mathbf{U}_s$.

Since the total momentum is conserved for any collision process, it follows that the collision term $\sum_s \delta_c \mathbf{p}_s / \delta t$ must sum to zero. In order to simplify the remaining terms, it is convenient to introduce a new pressure tensor

$$\overset{\leftrightarrow}{\mathbf{P}}_{0s} = m_s \int_{-\infty}^{\infty} (\mathbf{v} - \mathbf{U})(\mathbf{v} - \mathbf{U}) f_s d^3 v, \quad (6.1.6)$$

that gives the pressure in a frame of reference moving at the fluid velocity. It is easily verified that $\overset{\leftrightarrow}{\mathbf{P}}_{0s}$ is related to $\overset{\leftrightarrow}{\mathbf{P}}_s$, defined earlier by equation (??), via the equation

$$\overset{\leftrightarrow}{\mathbf{P}}_{0s} = \overset{\leftrightarrow}{\mathbf{P}}_s + \rho_{ms} \mathbf{W}_s \mathbf{W}_s, \quad (6.1.7)$$

where $\mathbf{W}_s = \mathbf{U}_s - \mathbf{U}$ is the average velocity of the s th species relative to the fluid velocity,

otherwise known as the *diffusion velocity* for species s . Writing the total pressure as $\overset{\leftrightarrow}{\mathbf{P}}_0 = \sum_s \overset{\leftrightarrow}{\mathbf{P}}_{0s}$, the momentum equation (6.1.5) becomes

$$\begin{aligned} & \sum_s \frac{\partial}{\partial t} (\rho_{ms} \mathbf{U}_s) + \sum_s \nabla \cdot (\rho_{ms} \mathbf{U}_s \mathbf{U}_s) \\ &= \rho_q \mathbf{E} + \mathbf{J} \times \mathbf{B} - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_0 + \sum_s \nabla \cdot (\rho_{ms} \mathbf{W}_s \mathbf{W}_s). \end{aligned} \quad (6.1.8)$$

With the substitution $\mathbf{U}_s = \mathbf{W}_s + \mathbf{U}$, the left-hand side of this equation can be written

$$\begin{aligned} & \frac{\partial}{\partial t} (\rho_{ms} \mathbf{U}) + \frac{\partial}{\partial t} (\rho_{ms} \mathbf{W}_s) + \nabla \cdot (\rho_{ms} \mathbf{W}_s \mathbf{W}_s) + \nabla \cdot (\rho_{ms} \mathbf{W}_s \mathbf{U}) \\ &+ \nabla \cdot (\rho_{ms} \mathbf{U} \mathbf{W}_s) + \nabla \cdot (\rho_{ms} \mathbf{U} \mathbf{U}). \end{aligned} \quad (6.1.9)$$

Summing over all species and using the identity $\sum_s \rho_{ms} \mathbf{W}_s = 0$ (implied by the definition of \mathbf{W}_s), this expression simplifies to

$$\sum_s \frac{\partial}{\partial t} (\rho_{ms} \mathbf{U}) + \sum_s \nabla \cdot (\rho_{ms} \mathbf{U} \mathbf{U}) + \sum_s \nabla \cdot (\rho_{ms} \mathbf{W}_s \mathbf{W}_s). \quad (6.1.10)$$

Substituting this result into the left-hand side of equation (6.1.8), the momentum equation simplifies to

$$\frac{\partial}{\partial t} (\rho_m \mathbf{U}) + \nabla \cdot (\rho_m \mathbf{U} \mathbf{U}) = \rho_q \mathbf{E} + \mathbf{J} \times \mathbf{B} - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_0. \quad (6.1.11)$$

Using the continuity equation (6.1.4) and following the same procedure used to rewrite the left-hand side of equation (??) in the form given by equation (??), the left-hand side of the above equation can be rewritten as

$$\rho_m \left[\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla) \mathbf{U} \right] = \rho_q \mathbf{E} + \mathbf{J} \times \mathbf{B} - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_0, \quad (6.1.12)$$

where the quantity in the brackets on the left is the convective derivative, $d\mathbf{U}/dt$, see equation (??). This equation is called the *momentum equation* and has the simple interpretation that the rate of change of the momentum of a fluid element is equal to the sum of the electric field, magnetic field and pressure forces acting on that fluid element. In what follows, we will drop the subscript 0 on the plasma pressure tensor, $\overset{\leftrightarrow}{\mathbf{P}}_0$ with the understanding that the pressure is always defined relative to the fluid frame of reference.

It is important to note that no assumptions, other than the validity of the Boltzmann equation, have been used in the derivation of the mass continuity and momentum equations.

6.1.3 Generalized Ohm's law

Since a plasma is a conducting medium, it is necessary to determine how the current density, \mathbf{J} , depends on the electric field, \mathbf{E} . In the historical approach it was assumed that the

current density can be represented by a simple Ohm's law of the form, $\mathbf{J}' = \sigma \mathbf{E}'$, where σ is the conductivity, and \mathbf{J}' and \mathbf{E}' are the current density and electric field in the rest frame of the fluid. For non-relativistic velocities, the current density and the electric field transform from the rest frame of the fluid to the laboratory frame according to the relations $\mathbf{J}' = \mathbf{J}$ and $\mathbf{E}' = \mathbf{E} + \mathbf{U} \times \mathbf{B}$, respectively. In the laboratory frame of reference Ohm's law then becomes

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{U} \times \mathbf{B}). \quad (6.1.13)$$

Unfortunately, a completely rigorous derivation of this equation is not possible. As will be shown below, numerous assumptions must be made to derive the above equation, some of which are questionable. Nevertheless, Ohm's law is widely used in MHD calculations, and is a reasonable approximation if the collisional mean-free path is short compared with the characteristic length scale of the system.

In an attempt to derive Ohm's law, let us consider a two-component plasma consisting of electrons (e) and one species of positively charged ions (i). Let us assume further that the drag force on the s th species is proportional to the velocity difference between the s th and r th species and is given by the equation

$$\frac{\delta_c \mathbf{p}_s}{\delta t} = -n_s m_s v_{sr} (\mathbf{U}_s - \mathbf{U}_r), \quad (6.1.14)$$

where v_{sr} is the effective collision frequency between the two species. The above equation is simply a generalized version of the collisional drag force given by equation (??), and is valid only if the velocity difference $\mathbf{U}_s - \mathbf{U}_r$ is small compared to the thermal velocities of the two species.

Next consider the moment equations for the electrons and the ions. The momentum equation for the ions is given by

$$\begin{aligned} \frac{\partial}{\partial t} (m_i n_i \mathbf{U}_i) + \nabla \cdot (m_i n_i \mathbf{U}_i \mathbf{U}_i) &= n_i e [\mathbf{E} + \mathbf{U}_i \times \mathbf{B}] \\ &\quad - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_i - n_i m_i v_{ie} (\mathbf{U}_i - \mathbf{U}_e), \end{aligned} \quad (6.1.15)$$

and the momentum equation for the electrons is given by

$$\begin{aligned} \frac{\partial}{\partial t} (m_e n_e \mathbf{U}_e) + \nabla \cdot (m_e n_e \mathbf{U}_e \mathbf{U}_e) &= -n_e e [\mathbf{E} + \mathbf{U}_e \times \mathbf{B}] \\ &\quad - \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_e - n_e m_e v_{ei} (\mathbf{U}_e - \mathbf{U}_i). \end{aligned} \quad (6.1.16)$$

Since the current density is proportional to $(\mathbf{U}_i - \mathbf{U}_e)$, we start by multiplying the above equations by m_e and m_i , respectively, and then subtracting the two equations in order to generate a $(\mathbf{U}_i - \mathbf{U}_e)$ term, which gives

$$\begin{aligned}
& m_e m_i \left\{ \frac{\partial}{\partial t} [n(\mathbf{U}_i - \mathbf{U}_e)] + \nabla \cdot [n(\mathbf{U}_i \mathbf{U}_i - \mathbf{U}_e \mathbf{U}_e)] \right\} \\
& = e \rho_m \mathbf{E} + en(m_e \mathbf{U}_i + m_i \mathbf{U}_e) \times \mathbf{B} - m_e \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_i + m_i \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_e \\
& - m_e m_i n v_{ie} (\mathbf{U}_i - \mathbf{U}_e) + m_i m_e n v_{ei} (\mathbf{U}_e - \mathbf{U}_i), \tag{6.1.17}
\end{aligned}$$

where $\rho_m = n(m_e + m_i)$ is the mass density and we have assumed that $n_e = n_i = n$. Since the drag force on the ions must be equal and opposite to the drag force on the electrons, it follows from equation (6.1.14) that $m_i v_{ie} = m_e v_{ei}$. The last two terms in equation (6.1.17) then simplify to

$$-m_e m_i n v_{ie} (\mathbf{U}_i - \mathbf{U}_e) + m_i m_e n v_{ei} (\mathbf{U}_e - \mathbf{U}_i) = -e \rho_m \left(\frac{m_e v_{ei}}{ne^2} \right) \mathbf{J}, \tag{6.1.18}$$

where $\mathbf{J} = en(\mathbf{U}_i - \mathbf{U}_e)$ is the current density. Comparing the term in the parentheses on the right-hand side of the above equation with equation (??) one can see that this term is simply the inverse of the conductivity

$$\frac{m_e v_{ei}}{ne^2} = \frac{1}{\sigma} = \eta, \tag{6.1.19}$$

which is called the resistivity. Next, consider the $m_e \mathbf{U}_i + m_i \mathbf{U}_e$ term on the right-hand side of equation (6.1.17). By simple algebraic manipulation of the equations for the fluid velocity, $\mathbf{U} = (m_e \mathbf{U}_e + m_i \mathbf{U}_i)/(m_e + m_i)$, and the current density, $\mathbf{J} = ne(\mathbf{U}_i - \mathbf{U}_e)$, it is easy to show that this term can be written in the form

$$m_e \mathbf{U}_i + m_i \mathbf{U}_e = (m_e + m_i) \mathbf{U} - (m_i - m_e) \frac{\mathbf{J}}{en}. \tag{6.1.20}$$

Finally, consider the left-hand side of equation (6.1.17). Using the equation for the current density, $\mathbf{J} = en(\mathbf{U}_i - \mathbf{U}_e)$, it is easy to see that the $\partial/\partial t$ term can be written

$$\frac{\partial}{\partial t} [n(\mathbf{U}_i - \mathbf{U}_e)] = \frac{1}{e} \frac{\partial \mathbf{J}}{\partial t}. \tag{6.1.21}$$

The ∇ term is more complicated, but can be simplified by noting that equation (6.1.14) is valid only if the velocity difference $\delta \mathbf{U} = \mathbf{U}_e - \mathbf{U}_i$ is small. The $(\mathbf{U}_i \mathbf{U}_i - \mathbf{U}_e \mathbf{U}_e)$ term can then be expanded in powers of the small quantity $\delta \mathbf{U}$ to give, to a first approximation,

$$\nabla \cdot [n(\mathbf{U}_i \mathbf{U}_i - \mathbf{U}_e \mathbf{U}_e)] = \frac{1}{e} \nabla \cdot (\mathbf{J} \mathbf{U} + \mathbf{U} \mathbf{J}), \tag{6.1.22}$$

where we have made use of the fact that $\mathbf{U} \approx \mathbf{U}_i$ for $m_e \ll m_i$, and $\mathbf{J} = -en\delta \mathbf{U}$. Substituting equations (6.1.18)–(6.1.22) into equation (6.1.17), making use of the fact that $m_e \ll m_i$, and omitting the ion pressure term since it is reduced by a factor of m_e/m_i relative to the electron pressure term, we obtain the equation

$$\mathbf{E} + \mathbf{U} \times \mathbf{B} - \frac{\mathbf{J}}{\sigma} = \frac{1}{en} \mathbf{J} \times \mathbf{B} - \frac{1}{en} \nabla \cdot \overset{\leftrightarrow}{\mathbf{P}}_e + \frac{m_e}{ne^2} \left[\frac{\partial \mathbf{J}}{\partial t} + \nabla \cdot (\mathbf{JU} + \mathbf{UJ}) \right]. \quad (6.1.23)$$

This equation is called the *generalized Ohm's law*. Comparing with equation (6.1.13), it can be seen that if all of the terms on the right-hand side are sufficiently small, the equation reduces to the simple form of Ohm's law. The last term on the left-hand side of the above equation as well as the terms on the right are often important in boundary layers where the current density, \mathbf{J} , is large. However, away from such regions, these terms can usually be neglected. A more detailed discussion of the validity of Ohm's law is given in Section 6.8.2.

As the collision frequency goes to zero, the conductivity σ goes to infinity. Ohm's law then reduces to $\mathbf{E} + \mathbf{U} \times \mathbf{B} = 0$. A plasma that obeys this equation is called an *ideal MHD plasma*. The fluid velocity component perpendicular to \mathbf{B} is then given by $\mathbf{U}_\perp = \mathbf{E} \times \mathbf{B}/B^2$, which is identical to the $\mathbf{E} \times \mathbf{B}$ drift velocity encountered in single particle orbit theory, see equation (??). Also, note that for an ideal MHD plasma, the electric field component parallel to the magnetic field, $E_{||}$, is identically zero. The magnetic field lines are then equipotentials (i.e., all points along a given magnetic field line are at the same potential) if $\partial \mathbf{A} / \partial t = 0$.

In collisional plasmas, it is often assumed that all of the terms on the right-hand side of equation (6.1.23) are negligible. This leads to the resistive form of Ohm's law

$$\mathbf{E} + \mathbf{U} \times \mathbf{B} = \frac{1}{\sigma} \mathbf{J} = \eta \mathbf{J}, \quad (6.1.24)$$

which is widely used, not because it is a rigorous model, but because it represents the simplest and qualitatively most important deviation from the ideal MHD model.

6.1.4 The equation of state

Since the moment equations do not define a closed system of equations, we must choose an equation of state in order to close the system of equations. The equation of state specifies the plasma pressure as a function of the temperature and density, and its form depends on various assumptions that must be made concerning the effect of collisions.

If there are sufficient collisions to establish an isotropic Maxwellian velocity distribution, then the plasma pressure is isotropic. The pressure gradient force can then be represented by the gradient of a scalar pressure, $\nabla \cdot \overset{\leftrightarrow}{\mathbf{P}} = \nabla P$, where we have dropped the subscript 0 on the pressure, which is now understood to be defined by equation (6.1.6). Under these conditions, the equation of state is commonly assumed to be a power law of the form

$$\frac{d}{dt}(P\rho_m^{-\gamma}) = 0, \quad (6.1.25)$$

where the exponent γ is called the polytrope index [Fetter and Walecka, 1980]. By choosing various values for γ , a variety of situations can be represented. For example, if the

compression is so rapid that no heat can flow, i.e., $dQ = 0$, then following the usual thermodynamic derivation, it can be shown that $\gamma = C_p/C_v$, where C_p is the heat capacity at constant pressure, and C_v is the heat capacity at constant volume. Equation (6.1.25) is then called the adiabatic equation of state. Note that the word “adiabatic” in this context means no heat flow, which is different from the meaning of adiabatic used in Chapter ??, which means slowly varying. Note also that, although the compression must be rapid, as in a sound wave, it must not be so rapid that the process is irreversible (i.e., not a succession of equilibrium states). This difficulty will be encountered later when we discuss shock waves. In statistical mechanics it can be shown that for the adiabatic equation of state the polytrope index is given by $\gamma = (f + 2)/f$, where f is the number of degrees of freedom. An MHD flow that is adiabatic is also sometimes called an isoentropic flow, since the entropy, $S = \int dQ/T$, is constant along the streamlines. Other equations of state that can be represented by equation (6.1.25) include an isothermal equation of state ($\gamma = 1$), and an incompressible equation of state ($\gamma = \infty$).

If there are insufficient collisions to maintain an isotropic velocity distribution, then the pressure is anisotropic and must be represented by a tensor. If the particle motions in the rest frame of the fluid are azimuthally symmetric with respect to the magnetic field, the pressure tensor can often be represented as a diagonal matrix of the form

$$\overset{\leftrightarrow}{\mathbf{P}} = \begin{pmatrix} P_{\perp} & 0 & 0 \\ 0 & P_{\perp} & 0 \\ 0 & 0 & P_{\perp} \end{pmatrix} \quad (6.1.26)$$

where P_{\perp} and P_{\parallel} are pressures perpendicular and parallel to the local magnetic field. For a slowly varying adiabatic process in which there is no heat flow, it can be shown that the parallel and perpendicular pressures obey the relations

$$\frac{d}{dt} \left(\frac{P_{\perp}}{\rho_m B} \right) = 0 \quad \text{and} \quad \frac{d}{dt} \left(\frac{P_{\perp}^2 P_{\parallel}}{\rho_m^5} \right) = 0. \quad (6.1.27)$$

The above system of equations is known as the Chew, Goldberger, and Low (CGL) equation of state (see Section 5.4.4.3). As previously discussed, the first equation is equivalent to conservation of the 1st adiabatic invariant (see Section 3.6.2.1), and the second equation is equivalent to conservation of the 2nd adiabatic invariant (see Section 3.6.2.2).

6.1.5 The complete set of resistive MHD equations

In order to obtain a complete and self-consistent description of a resistive MHD plasma, the fluid equations and Ohm's law must be combined with Maxwell's equations. Two assumptions are made concerning the detailed form of Maxwell's equations. First, since temporal variations are assumed to be slow, the displacement current, $\epsilon_0 \partial \mathbf{E} / \partial t$, is ignored in comparison with the conduction current. Second, since the fluid is assumed to be electrically neutral, the charge density is taken to be identically zero, $\rho_q = 0$. In addition, in resistive MHD we assume that the pressure is a scalar and obeys equation (6.1.25). With these assumptions, the complete set of resistive MHD equations becomes

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \quad (\text{Ampère's law}), \quad (6.1.28)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (6.1.29)$$

$$\nabla \times \mathbf{E} = -\partial \mathbf{B} / \partial t \quad (\text{Faraday's law}), \quad (6.1.30)$$

$$\nabla \cdot \mathbf{E} = 0 \quad (\text{Gauss' law}), \quad (6.1.31)$$

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \mathbf{U}) = 0 \quad (\text{mass continuity equation}), \quad (6.1.32)$$

$$\rho_m \frac{d\mathbf{U}}{dt} = \mathbf{J} \times \mathbf{B} - \nabla P \quad (\text{momentum equation}), \quad (6.1.33)$$

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{U} \times \mathbf{B}) \quad (\text{Ohm's law}), \quad (6.1.34)$$

and

$$\frac{d}{dt}(P\rho_m^{-\gamma}) = 0 \quad (\text{equation of state}). \quad (6.1.35)$$

If we take the divergence of Ampère's law (6.1.28), we obtain the condition for charge continuity, $\nabla \cdot \mathbf{J} = 0$, which is always valid when it is assumed that $\rho_q = 0$. Note that the charge continuity equation is not an independent equation, since it is implied by Ampère's law.

6.2 Magnetic pressure

Further insight into the physical nature of the interaction of the magnetic field with an MHD plasma can be obtained by using Ampère's law (6.1.28) to eliminate \mathbf{J} from the $\mathbf{J} \times \mathbf{B}$ force term in the momentum equation, which then becomes

$$\mathbf{J} \times \mathbf{B} = \frac{1}{\mu_0}(\nabla \times \mathbf{B}) \times \mathbf{B}. \quad (6.2.1)$$

Using the identity $\nabla(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{F} \cdot \nabla)\mathbf{G} + \mathbf{F} \times (\nabla \times \mathbf{G}) + (\mathbf{G} \cdot \nabla)\mathbf{F} + \mathbf{G} \times (\nabla \times \mathbf{F})$ with $\mathbf{F} = \mathbf{G} = \mathbf{B}$, the $\mathbf{J} \times \mathbf{B}$ force can be rewritten in the form

$$\mathbf{J} \times \mathbf{B} = \frac{1}{\mu_0}(\mathbf{B} \cdot \nabla)\mathbf{B} - \nabla \left(\frac{B^2}{2\mu_0} \right) \quad (6.2.2)$$

Since the pressure term in the momentum equation involves the negative of the divergence of the pressure tensor, it is interesting to ask whether the terms on the right-hand side of the above equation can be expressed as the negative of the divergence of a tensor. The answer is that they can. The appropriate tensor $\overset{\leftrightarrow}{T}$ is given in index notation by

$$T_{rs} = -\frac{B_r B_s}{\mu_0} + \delta_{rs} \frac{B^2}{2\mu_0}, \quad (6.2.3)$$

where δ_{rs} is the Kronecker delta ($\delta_{rs} = 1$ for $r = s$, $\delta_{rs} = 0$ for $r \neq s$), or in vector notation by

$$\overset{\leftrightarrow}{\mathbf{T}} = -\frac{\mathbf{B}\mathbf{B}}{\mu_0} + \overset{\leftrightarrow}{\mathbf{1}} = \frac{B^2}{2\mu_0}. \quad (6.2.4)$$

To show this relationship, consider the s component of the negative of the divergence of this tensor,

$$\begin{aligned} -(\nabla \cdot \overset{\leftrightarrow}{\mathbf{T}})_s &= -\sum_r \frac{\partial}{\partial x_r} \left[-\frac{B_r B_s}{\mu_0} + \delta_{rs} \frac{B^2}{2\mu_0} \right] \\ &= \frac{1}{\mu_0} \sum_r \left[B_r \frac{\partial B_s}{\partial x_r} + B_s \frac{\partial B_r}{\partial x_r} \right] - \frac{\partial}{\partial x_s} \frac{B^2}{2\mu_0}. \end{aligned} \quad (6.2.5)$$

Since $\nabla \cdot \mathbf{B} = 0$, which in index notation is $\sum_r (\partial B_r / \partial x_r) = 0$, it is easy to see that the above equation simplifies to

$$-(\nabla \cdot \overset{\leftrightarrow}{\mathbf{T}})_s = \frac{1}{\mu_0} \sum_r B_r \frac{\partial B_s}{\partial x_r} - \frac{\partial}{\partial x_s} \left(\frac{B^2}{2\mu_0} \right), \quad (6.2.6)$$

which in vector notation can be written

$$-\nabla \cdot \overset{\leftrightarrow}{\mathbf{T}} = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \nabla \left(\frac{B^2}{2\mu_0} \right). \quad (6.2.7)$$

Comparing the right-hand side of this equation with the right-hand side of equation (6.2.2), we see that, as advertised, $\mathbf{J} \times \mathbf{B} = -\nabla \cdot \overset{\leftrightarrow}{\mathbf{T}}$. The tensor $\overset{\leftrightarrow}{\mathbf{T}}$ is called the magnetic pressure tensor. Using the magnetic pressure tensor, the momentum equation (6.1.33) can then be written in the compact form

$$\rho_m \frac{d\mathbf{U}}{dt} = -\nabla \cdot (\overset{\leftrightarrow}{\mathbf{T}} + \overset{\leftrightarrow}{\mathbf{P}}). \quad (6.2.8)$$

From the structure of the terms on the right-hand side of the above equation, it is apparent that the magnetic field produces an anisotropic “magnetic pressure” that simply adds to the plasma pressure. The anisotropic nature of the magnetic pressure can be illustrated by choosing a coordinate system in which the magnetic field is parallel to the z axis, $\mathbf{B} = (0, 0, B)$. In this coordinate system the magnetic pressure tensor can be written as the sum of two terms

$$\overset{\leftrightarrow}{\mathbf{T}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -B^2/\mu_0 \end{bmatrix} + \begin{bmatrix} B^2/2\mu_0 & 0 & 0 \\ 0 & B^2/2\mu_0 & 0 \\ 0 & 0 & B^2/\mu_0 \end{bmatrix}. \quad (6.2.9)$$

The above equation shows that the magnetic pressure is the sum of a negative pressure (i.e., a *tension*) B^2/μ_0 that acts along the magnetic field, plus a positive pressure $B^2/2\mu_0$ that acts equally in all directions.

Various other forms of the momentum equation are often useful. If the plasma pressure is isotropic, as is usually assumed in resistive MHD, the momentum equation can be written

$$\rho_m \frac{d\mathbf{U}}{dt} = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \nabla \left(\frac{B^2}{2\mu_0} + P \right), \quad (6.2.10)$$

which explicitly shows that the isotropic part of the magnetic pressure simply adds to the plasma pressure. In many applications it is useful to rewrite the $\rho_m d\mathbf{U}/dt$ term on the left-hand side of the above equation as $\partial(\rho_m \mathbf{U})/\partial t + \nabla \cdot (\rho_m \mathbf{U} \mathbf{U})$, in which case the momentum equation becomes

$$\frac{\partial}{\partial t} (\rho_m \mathbf{U}) + \nabla \cdot (\rho_m \mathbf{U} \mathbf{U}) = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \nabla \left(\frac{B^2}{2\mu_0} + P \right). \quad (6.2.11)$$

It often occurs in the above equations that one or other of the two terms, $B^2/2\mu_0$ or P , dominates, in which case the smaller of the two terms can be neglected. To characterize the relative importance of the two terms, it is useful to define a quantity called the plasma β , which is given by

$$\beta = \frac{P}{(B^2/2\mu_0)}. \quad (6.2.12)$$

If $\beta \gg 1$, the plasma pressure force dominates, and the magnetic field pressure can be neglected. On the other hand, if $\beta \ll 1$, the magnetic field force dominates and the plasma pressure can be neglected.

6.3 Magnetic field convection and diffusion

By a relatively simple series of manipulations, the resistive MHD equations can be rearranged in a form that gives information on the time evolution of the magnetic field. Eliminating \mathbf{J} between Ampère's law (6.1.28) and Ohm's law (6.1.34), gives

$$\nabla \times \mathbf{B} = \mu_0 \sigma (\mathbf{E} + \mathbf{U} \times \mathbf{B}). \quad (6.3.1)$$

Taking the curl of the above equation,

$$\nabla \times (\nabla \times \mathbf{B}) = \mu_0 \sigma (\nabla \times \mathbf{E} + \nabla \times (\mathbf{U} \times \mathbf{B})), \quad (6.3.2)$$

using the identity

$$\nabla \times (\nabla \times \mathbf{B}) = \nabla (\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B}, \quad (6.3.3)$$

with $\nabla \cdot \mathbf{B} = 0$, and using Faraday's law (6.1.30), then gives the equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B}) + \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B}. \quad (6.3.4)$$

The above equation shows that the time rate of change of the magnetic field is controlled by two terms. The first term, which involves the fluid velocity, is called the *convection term*, and the second term, which involves the conductivity, is called the *diffusion term*. Usually for a time dependent magnetic field, one or other of these two terms dominates. To determine which term dominates, it is useful to introduce a dimensionless parameter R_m , called the magnetic Reynolds number, which is defined as the ratio of the typical magnitude of the convection term to the typical magnitude of the diffusion term,

$$R_m = \frac{|\nabla \times (\mathbf{U} \times \mathbf{B})|}{\left| \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B} \right|}. \quad (6.3.5)$$

The magnetic Reynolds number is analogous to the Reynolds number commonly used in fluid mechanics [see Batchelor, 1967]. To make a rough estimate of the magnetic Reynolds number, replace ∇ by $1/L$, where L is a length scale that characterizes the spatial gradients in the fluid, and ignore the vector character of the equation. The ratio of the magnitude of the convection term to the magnitude of the diffusion term is then given by

$$R_m = \mu_0 \sigma U L. \quad (6.3.6)$$

If $R_m \gg 1$, the convection term dominates, and if $R_m \ll 1$, the diffusion term dominates.

6.3.1 $R_m \gg 1$, the frozen field theorem

Large magnetic Reynolds numbers occur whenever the conductivity, the fluid velocity, and the length scale are sufficiently large to make the product $\mu_0 \sigma_0 U L \gg 1$. The diffusion term is then much smaller than the convection term and, to a first approximation, can be ignored. When $R_m \gg 1$ equation (6.3.4) simplifies to

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B}). \quad (6.3.7)$$

The above equation can be used to prove an important result known as the “frozen field theorem.” Two different, but equivalent, statements of the frozen field theorem can be made, each of which involves a somewhat different proof.

Statement 1. The magnetic flux threading any closed curve moving with the fluid is constant.

To prove Statement 1, consider the magnetic flux through a closed curve C at time t , shown in Figure 6.1. Suppose that curve C moves with the fluid, with each point on the curve moving to a new point $\mathbf{U} dt$ after a time dt . It is then easy to see that the rate of change of the magnetic flux through curve C is given by the sum of two terms

$$\frac{d\Phi_B}{dt} = \int_s \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{A} + \int_C \mathbf{B} \cdot (\mathbf{U} \times d\ell). \quad (6.3.8)$$

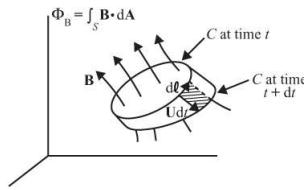


Fig. 6.1 The temporal evolution of the magnetic flux through a curve C moving with the fluid from time t to time $t + dt$.

The first term gives the rate of change of the flux due to the explicit time dependence of \mathbf{B} , and the second term gives the rate of change of the flux due to the motion of the curve C . The triple vector product in the second term can be permuted once to give

$$\frac{d\Phi_B}{dt} = \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{A} + \int_C d\ell \cdot (\mathbf{B} \times \mathbf{U}). \quad (6.3.9)$$

By Stokes' theorem and equation (6.3.7), it is easy to see that

$$\frac{d\Phi_B}{dt} = \int_S \left[\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{U} \times \mathbf{B}) \right] \cdot d\mathbf{A} = 0, \quad (6.3.10)$$

which proves Statement 1. This theorem has the following consequence. Consider two curves, C_1 and C_2 that are connected by magnetic field lines at time t , as shown in Figure 6.2. These field lines form a flux tube with a total magnetic flux Φ_B . As the fluid moves, since $d\Phi_B/dt = 0$ on the surfaces that make up the tube, it is easy to see that the tube continues to contain the same total magnetic flux at any later time t' . Since the ends of the flux tube can be shrunk to infinitesimal size, the flux tube becomes a magnetic field line, which then justifies the statement that “the magnetic field lines are frozen into the fluid.”

Statement 2. If a line moving with the fluid is a magnetic field line initially, it will be so for all times.

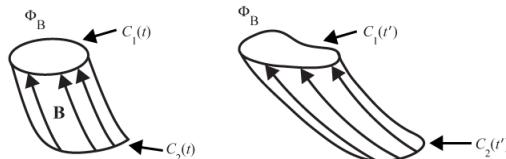


Fig. 6.2 For large magnetic Reynolds numbers, the magnetic flux, Φ_B , is “frozen” into the fluid.

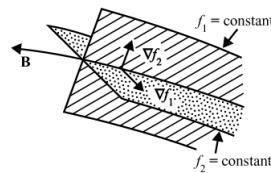


Fig. 6.3 Surfaces of constant f_1 and f_2 define a magnetic field line.

To prove Statement 2, we start by introducing the concept of *flux coordinates*. Consider the partial differential equation

$$\mathbf{B} \cdot \nabla f = 0. \quad (6.3.11)$$

Since ∇f is perpendicular to a contour of constant f , it follows that f is a constant along a magnetic field line. Consider two functionally independent families of surfaces, f_1 and f_2 , that obey

$$\mathbf{B} \cdot \nabla f_1 = \mathbf{B} \cdot \nabla f_2 = 0. \quad (6.3.12)$$

A magnetic field line can then be specified by the intersection of two such surfaces, as shown in Figure 6.3. Since \mathbf{B} is perpendicular to ∇f_1 and to ∇f_2 , we can think of \mathbf{B} , ∇f_1 , and ∇f_2 as a triad of vectors in three-dimensional space, such that

$$\mathbf{B} = \nabla f_1 \times \nabla f_2, \quad (6.3.13)$$

provided f_1 and f_2 are suitably normalized. The variables f_1 and f_2 are known as flux coordinates. By computing the divergence of equation (6.3.13), it is easy to show that this representation automatically satisfies the divergence-free condition $\nabla \cdot \mathbf{B} = 0$.

With this preparation, we are now ready to proceed with the proof of Statement 2. The basic approach is to compute the time rate of change of $\mathbf{B} \cdot \nabla f$ along a representative fluid trajectory. If we can show that the rate of change is identically zero, then Statement 2 is proven. To compute the time rate of change, we must use the convective derivative for d/dt , so that

$$\frac{d}{dt}(\mathbf{B} \cdot \nabla f) = \frac{\partial}{\partial t}(\mathbf{B} \cdot \nabla f) + (\mathbf{U} \cdot \nabla)(\mathbf{B} \cdot \nabla f). \quad (6.3.14)$$

The first term of the above equation can be expanded as follows,

$$\frac{\partial}{\partial t}(\mathbf{B} \cdot \nabla f) = \frac{\partial \mathbf{B}}{\partial t} \cdot \nabla f + \mathbf{B} \cdot \nabla \frac{\partial f}{\partial t}. \quad (6.3.15)$$

Using equation (6.3.7), the $\partial \mathbf{B} / \partial t$ term in the above equation can be written in the form

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla)\mathbf{U} - (\nabla \cdot \mathbf{U})\mathbf{B} - (\mathbf{U} \cdot \nabla)\mathbf{B}, \quad (6.3.16)$$

where we have used the rule for the curl of a cross-product, $\nabla \times (\mathbf{F} \times \mathbf{G}) = \mathbf{F}(\nabla \cdot \mathbf{G}) - \mathbf{G}(\nabla \cdot \mathbf{F}) + (\mathbf{G} \cdot \nabla)\mathbf{F} - (\mathbf{F} \cdot \nabla)\mathbf{G}$. To proceed further, we must require that the surfaces of constant f move with the fluid. This condition can be imposed by requiring that

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \mathbf{U} \cdot \nabla f = 0, \quad (6.3.17)$$

where $d/dt = \partial/\partial t + \mathbf{U} \cdot \nabla$ is the convective derivative.

Solving the above equation for $\partial f / \partial t$ and substituting the result into equation (6.3.15), and then substituting equations (6.3.15) and (6.3.16) into equation (6.3.14) gives

$$\frac{d}{dt}(\mathbf{B} \cdot \nabla f) = [(\mathbf{B} \cdot \nabla) \mathbf{U}] \cdot \nabla f - (\nabla \cdot \mathbf{U})(\mathbf{B} \cdot \nabla f) - [(\mathbf{U} \cdot \nabla) \mathbf{B}] \cdot \nabla f - \mathbf{B} \cdot \nabla [(\mathbf{U} \cdot \nabla) f] + (\mathbf{U} \cdot \nabla)(\mathbf{B} \cdot \nabla f). \quad (6.3.18)$$

The first, third, and fourth terms on the right-hand side of the above equation can be simplified by using index notation, as shown below,

$$\begin{aligned} &[(\mathbf{B} \cdot \nabla) \mathbf{U}] \cdot \nabla f - [(\mathbf{U} \cdot \nabla) \mathbf{B}] \cdot \nabla f - \mathbf{B} \cdot \nabla [(\mathbf{U} \cdot \nabla) f] \\ &= \sum_{i,j} \left[B_i \frac{\partial U_j}{\partial x_i} \frac{\partial f}{\partial x_j} - U_i \frac{\partial B_j}{\partial x_i} \frac{\partial f}{\partial x_j} - B_i \frac{\partial}{\partial x_i} \left(U_j \frac{\partial f}{\partial x_j} \right) \right] \\ &= \sum_{i,j} \left[B_i \frac{\partial U_j}{\partial x_i} \frac{\partial f}{\partial x_j} - U_i \frac{\partial B_j}{\partial x_i} \frac{\partial f}{\partial x_j} - B_i \frac{\partial U_j}{\partial x_i} \frac{\partial f}{\partial x_j} - B_i U_j \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} \right] \\ &= -(\mathbf{U} \cdot \nabla)(\mathbf{B} \cdot \nabla f). \end{aligned} \quad (6.3.19)$$

Note that the final expression in equation (6.3.19) is obtained by combining the second and last terms inside the parentheses, while the first and third terms cancel each other. Substituting the above expression into equation (6.3.18) gives the result that we are seeking, which is

$$\frac{d}{dt}(\mathbf{B} \cdot \nabla f) = -(\nabla \cdot \mathbf{U})(\mathbf{B} \cdot \nabla f). \quad (6.3.20)$$

The above equation shows that if $\mathbf{B} \cdot \nabla f = 0$ at $t = 0$, then $\mathbf{B} \cdot \nabla f = 0$ for all times. Hence, if the contours of constant f_1 and f_2 label a magnetic field line (i.e., $\mathbf{B} \cdot \nabla f = 0$) at $t = 0$, then they label a magnetic field line for all times, which proves Statement 2.



Fig. 6.4 Stretching a magnetic flux tube that is frozen into the fluid increases the magnetic field strength.

From the frozen field theorem, we can assert that, if at $t = 0$ we label a field line by tagging fluid particles on it, then the same fluid particles will remain frozen to the field line at all subsequent times. It follows that in such a plasma, field lines cannot intersect or break. For if they did, the velocity of the fluid particles at the break-point or the point of intersection would have to change discontinuously. In other words, the “topology” of the magnetic field lines cannot change.

An interesting consequence of the frozen field theorem is that the magnetic field strength can be amplified by changes in the fluid geometry or the mass density. Consider a fluid element that is stretched between two times t and t' , as shown in Figure 6.4. The frozen

field theorem shows that $B\Delta A = B'\Delta A'$. From mass conservation, one also has the relation $\rho_m \Delta A \ell = \rho'_m \Delta A' \ell'$. Eliminating ΔA and $\Delta A'$ between these two equations gives

$$B' = B \left(\frac{\rho'_m}{\rho_m} \right) \left(\frac{\ell'}{\ell} \right). \quad (6.3.21)$$

This equation shows that the magnetic field strength increases whenever either the mass density ρ'_m or the length ℓ' increases. An example of magnetic amplification occurs when a star collapses to form a neutron star. As the star collapses, the density ($\rho'_m \sim 1/\ell'^3$) increases much faster than the length ℓ' decreases, so the magnetic field is increased by a factor of $(\ell/\ell')^2$, which can be very large. Field strengths as large as 10^{11} to 10^{12} gauss can be produced at the surface of a neutron star from the ~ 1 gauss surface field of the parent star.

An example of magnetic amplification due to stretching occurs in the solar atmosphere. Because the Sun rotates faster near the equator than near the poles, magnetic field lines near the equator are stretched azimuthally around the Sun, as shown in Figure 6.5. Since the stretching motion is nearly horizontal, the density remains essentially constant ($\rho' \sim \rho$). As the length of the field line increases, the magnetic field strength increases ($B \sim \ell'$). Eventually the magnetic field becomes so strong that the Rayleigh-Taylor instability (see Section 6.7.7) causes the magnetic field lines to bulge up through the surface of the Sun [Elsasser, 1956; Babcock, 1961]. The strong magnetic field in the bulge inhibits the transport of heat from deeper layers in the Sun, causing sunspots, which are cooler and darker than the surrounding gas.



Fig. 6.5

Differential rotation at the Sun stretches the magnetic field lines, thereby increasing the magnetic field strength.

6.3.2 $R_m \ll 1$, magnetic diffusion

If the magnetic Reynolds number is much less than one, then the convection term is much smaller than the diffusion term, and to a first approximation can be ignored. Equation (6.3.4) then simplifies to

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B}. \quad (6.3.22)$$

The above equation has the mathematical form of a diffusion equation. Therefore, the magnetic field diffuses through the fluid in a manner analogous to the diffusion of heat through a solid. The characteristic time scale for the diffusion, T , can be estimated by substituting $1/T$ for $\partial/\partial t$ and $1/L^2$ for ∇^2 , which, ignoring the vector aspect of the equation, gives

Table 6.1 Diffusion times for various conducting systems

Conductor	T
Copper sphere (10 cm diameter)	10^{-1} s
Fusion machine	10 s
Earth's molten iron core	10^4 years
Interior of the Sun	10^{10} years

$$T = \mu_0 \sigma L^2. \quad (6.3.23)$$

Some representative diffusion times are given in Table 6.1. These examples show that for most laboratory applications the diffusion times are quite short, typically on the order of seconds. On the other hand, for most geophysical and astrophysical applications, the diffusion times are extremely long, sometimes even longer than the age of the Universe. Therefore, for most geophysical and astrophysical applications, the magnetic field can be considered to be “frozen” into the fluid (i.e., $R_m \gg 1$). However, even in such large-scale systems, the frozen field concept sometimes breaks down in small localized regions where the scale size becomes very small. Such effects are discussed in more detail in Section 6.8 under the topic of magnetic reconnection.

6.4 The energy equation

In the limit of ideal MHD (i.e., $\sigma \rightarrow \infty$), the MHD equations conserve energy exactly. To derive an expression for the energy in this limiting case, we start with the momentum equation (6.1.33). Multiplying the momentum equation by \mathbf{U} and eliminating \mathbf{J} via Ampère’s law (6.1.28) gives

$$\rho_m \mathbf{U} \cdot \left[\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla) \mathbf{U} \right] = -(\mathbf{U} \cdot \nabla) P + \frac{\mathbf{U}}{\mu_0} \cdot [(\nabla \times \mathbf{B}) \times \mathbf{B}]. \quad (6.4.1)$$

Using the mass continuity equation (6.1.32), the left-hand side of the above equation can be rewritten in the following form:

$$\begin{aligned}
& \rho_m \mathbf{U} \cdot \left[\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla) \mathbf{U} \right] \\
&= \frac{1}{2} \rho_m \left(\frac{\partial U^2}{\partial t} + \mathbf{U} \cdot \nabla U^2 \right) \\
&= \frac{\partial}{\partial t} \left(\frac{1}{2} \rho_m U^2 \right) - \frac{U^2}{2} \frac{\partial \rho_m}{\partial t} + \frac{1}{2} \rho_m \mathbf{U} \cdot \nabla U^2 \\
&= \frac{\partial}{\partial t} \left(\frac{1}{2} \rho_m U^2 \right) + \frac{U^2}{2} \nabla \cdot (\rho_m \mathbf{U}) + \frac{1}{2} \rho_m \mathbf{U} \cdot \nabla U^2 \\
&= \frac{\partial}{\partial t} \left(\frac{1}{2} \rho_m U^2 \right) + \nabla \cdot \left(\frac{U^2}{2} \rho_m \mathbf{U} \right).
\end{aligned} \tag{6.4.2}$$

Next, consider the $(\mathbf{U} \cdot \nabla)P$ term on the right-hand side of equation (6.4.1). Since this term involves the pressure, for an equation of state we use the adiabatic equation of state given by equation (6.1.35) which, when expanded, becomes

$$\frac{d}{dt} \left(\frac{P}{\rho_m^\gamma} \right) = \left(\frac{dP}{dt} \right) \rho_m^{-\gamma} - \gamma P \rho_m^{-(\gamma+1)} \frac{d\rho_m}{dt} = 0. \tag{6.4.3}$$

The above equation simplifies to

$$\frac{dP}{dt} = \frac{\gamma P}{\rho_m} \frac{d\rho_m}{dt}. \tag{6.4.4}$$

To proceed further, the total derivatives in this equation must be replaced, using the convective derivative ($d/dt = \partial/\partial t + \mathbf{U} \cdot \nabla$), which gives

$$\frac{\partial P}{\partial t} + (\mathbf{U} \cdot \nabla)P = \frac{\gamma P}{\rho_m} \left[\frac{\partial \rho_m}{\partial t} + (\mathbf{U} \cdot \nabla) \rho_m \right]. \tag{6.4.5}$$

Note that the desired term, $(\mathbf{U} \cdot \nabla)P$, has now appeared on the left-hand side of the equation. Using the mass continuity equation (6.1.32) to replace the $\partial \rho_m / \partial t$ term, the above equation simplifies to

$$\frac{\partial P}{\partial t} + (\mathbf{U} \cdot \nabla)P = -\gamma P \nabla \cdot \mathbf{U}. \tag{6.4.6}$$

Using the identity $P \nabla \cdot \mathbf{U} = \nabla \cdot (P \mathbf{U}) - \mathbf{U} \cdot \nabla P$, the above equation can be written

$$\frac{\partial P}{\partial t} + (\mathbf{U} \cdot \nabla)P = -\gamma \nabla \cdot (P \mathbf{U}) + \gamma \mathbf{U} \cdot \nabla P, \tag{6.4.7}$$

which can be solved for $(\mathbf{U} \cdot \nabla)P$ to give

$$(\mathbf{U} \cdot \nabla)P = \frac{1}{\gamma - 1} \frac{\partial P}{\partial t} + \frac{\gamma}{\gamma - 1} \nabla \cdot (P \mathbf{U}). \tag{6.4.8}$$

Returning again to equation (6.4.1), the second term on the right can be rewritten as

$$\mathbf{U} \cdot (\nabla \times \mathbf{B}) \times \mathbf{B} = -(\mathbf{U} \times \mathbf{B}) \cdot (\nabla \times \mathbf{B}). \quad (6.4.9)$$

Noting that $\mathbf{E} = -\mathbf{U} \times \mathbf{B}$ for an ideal MHD fluid, the term on the right can be expanded using the identity $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G})$ to give

$$\begin{aligned} \mathbf{U} \cdot (\nabla \times \mathbf{B}) \times \mathbf{B} &= \mathbf{E} \cdot (\nabla \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \nabla \cdot (\mathbf{E} \times \mathbf{B}) \\ &= -\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} - \nabla \cdot (\mathbf{E} \times \mathbf{B}) \\ &= -\frac{\partial}{\partial t} \left(\frac{B^2}{2} \right) - \nabla \cdot (\mathbf{E} \times \mathbf{B}). \end{aligned} \quad (6.4.10)$$

Finally, substituting equations (6.4.2), (6.4.8), and (6.4.10) into equation (6.4.1), we obtain the energy conservation law for an ideal MHD fluid

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho_m U^2 + \frac{P}{\gamma - 1} + \frac{B^2}{2\mu_0} \right) + \nabla \cdot \left(\frac{1}{2} \rho_m U^2 \mathbf{U} + \frac{\gamma}{\gamma - 1} P \mathbf{U} + \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right) = 0. \quad (6.4.11)$$

In thermodynamics, the term

$$\xi = P/(\gamma - 1) \quad (6.4.12)$$

is called the *internal energy* and the term

$$h = \gamma P/(\gamma - 1) \quad (6.4.13)$$

is called the *enthalpy*. Although equation (6.4.11) was derived using the adiabatic equation of state, which is a reversible process, equations (6.4.12) and (6.4.13) for the internal energy and enthalpy are also applicable to irreversible processes, provided we use $\gamma = C_p/C_v$, or equivalently $\gamma = (f+2)/f$. The reason these equations are valid for irreversible as well as reversible processes is that they are functions of the state of the system and do not depend on the detailed process by which this state is achieved. Other terms recognizable in equation (6.4.11) are the magnetic energy density, $B^2/2\mu_0$, the kinetic energy density, $(1/2)\rho_m U^2$, and the Poynting flux, $\mathbf{S} = (1/\mu_0)\mathbf{E} \times \mathbf{B}$.

If we integrate the energy conservation law over all space, the divergence term can be converted to a surface integral at infinity, which gives

$$\frac{\partial}{\partial t} \int_V \left(\frac{1}{2} \rho_m U^2 + \xi + \frac{B^2}{2\mu_0} \right) d^3x + \int_S \left(\frac{1}{2} \rho_m U^2 \mathbf{U} + h \mathbf{U} + \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right) \cdot d\mathbf{A} = 0, \quad (6.4.14)$$

where we have introduced the symbols ξ and h for the internal energy and the enthalpy, respectively. Assuming that the surface integral vanishes at infinity, the above equation simplifies to

$$\frac{\partial}{\partial t} \int_V \left(\frac{1}{2} \rho_m U^2 + \xi + \frac{B^2}{2\mu_0} \right) d^3x = 0, \quad (6.4.15)$$

which can be written in the form of an energy conservation equation

$$K + W = \text{constant}, \quad (6.4.16)$$

where

$$K = \int_V \frac{1}{2} \rho_m U^2 d^3x \quad (6.4.17)$$

is the kinetic energy, and

$$W = \int_V \left(\xi + \frac{B^2}{2\mu_0} \right) d^3x \quad (6.4.18)$$

is the potential energy. These energy conservation equations will later be useful when we consider the stability of various MHD systems.

6.5 Magnetohydrodynamic waves

Next we consider the propagation of *small-amplitude* waves in a homogeneous ideal (infinitely conducting) MHD fluid. The relevant equations are Faraday's law (6.1.30), the mass continuity equation (6.1.32), the momentum equation (6.1.33), and an equation of state, which we take to be the adiabatic equation of state (6.1.35), since waves usually occur on time scales sufficiently short that no heat flows. These equations are linearized in the usual way by assuming that \mathbf{U} , ρ_m , \mathbf{B} , and P are the sum of a spatially uniform time independent zero-order (0) quantity plus a small first-order (1) perturbation, i.e., $\mathbf{U} = \mathbf{U}_1$, $\rho_m = \rho_{m0} + \rho_{m1}$, $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1$, and $P = P_0 + P_1$. The linearized *first-order* equations of motion are easily shown to be

$$\frac{\partial \rho_{m1}}{\partial t} + \rho_{m0} \nabla \cdot \mathbf{U}_1 = 0 \quad (6.5.1)$$

$$\rho_{m0} \frac{\partial \mathbf{U}_1}{\partial t} = \frac{1}{\mu_0} (\nabla \times \mathbf{B}_1) \times \mathbf{B}_0 - \nabla P_1 \quad (6.5.2)$$

$$\frac{\partial \mathbf{B}_1}{\partial t} = \nabla \times (\mathbf{U}_1 \times \mathbf{B}_0) \quad (6.5.3)$$

and

$$P_1 = \gamma \left(\frac{P_0}{\rho_{m0}} \right) \rho_{m1}. \quad (6.5.4)$$

In the adiabatic equation of state (6.5.4), it is convenient to introduce the speed of sound V_S , which is defined by the equation

$$V_S^2 = \gamma \frac{P_0}{\rho_{m0}} = \frac{\gamma \kappa T_0}{m}, \quad (6.5.5)$$

where T_0 is the zero-order temperature, and m is the average molecular mass. Next, we Fourier analyze the above equations by making the usual operator substitutions, $\nabla \rightarrow i\mathbf{k}$, and $\partial/\partial t \rightarrow -i\omega$. At this point we also drop the subscript 1 on the first-order terms, which can be distinguished from the zero-order terms by the subscript 0 on the zero-order terms. The resulting equations are

$$-i\omega \tilde{\rho}_m + i\rho_{m0} \mathbf{k} \cdot \tilde{\mathbf{U}} = 0 \quad (6.5.6)$$

$$-i\omega \rho_{m0} \tilde{\mathbf{U}} = \frac{i}{\mu_0} (\mathbf{k} \times \tilde{\mathbf{B}}) \times \mathbf{B}_0 - i\mathbf{k} \tilde{P} \quad (6.5.7)$$

$$-i\omega \tilde{\mathbf{B}} = i\mathbf{k} \times (\tilde{\mathbf{U}} \times \mathbf{B}_0) \quad (6.5.8)$$

$$\tilde{P} = V_S^2 \tilde{\rho}_m. \quad (6.5.9)$$

Eliminating $\tilde{\rho}_m$ from equation (6.5.9) by using equation (6.5.6) we obtain an equation for the pressure

$$\tilde{P} = V_S^2 \frac{\rho_{m0}}{\omega} \mathbf{k} \cdot \tilde{\mathbf{U}}. \quad (6.5.10)$$

Using the above equation, the pressure \tilde{P} can be eliminated from equation (6.5.7), which, after multiplying by $i\omega/\rho_{m0}$, becomes

$$\omega^2 \tilde{\mathbf{U}} = \frac{-\omega}{\mu_0 \rho_{m0}} (\mathbf{k} \times \tilde{\mathbf{B}}) \times \mathbf{B}_0 + V_S^2 \mathbf{k} (\mathbf{k} \cdot \tilde{\mathbf{U}}) \quad (6.5.11)$$

Finally, $\tilde{\mathbf{B}}$ can be eliminated from the above equation by using equation (6.5.8), which gives the homogeneous equation for the fluid velocity

$$\omega^2 \tilde{\mathbf{U}} = \frac{1}{\mu_0 \rho_{m0}} \{ \mathbf{k} \times (\mathbf{k} \times [\tilde{\mathbf{U}} \times \mathbf{B}_0]) \} \times \mathbf{B}_0 + V_S^2 \mathbf{k} (\mathbf{k} \cdot \tilde{\mathbf{U}}). \quad (6.5.12)$$

With no loss in generality we can assume that $\mathbf{B}_0 = (0, 0, B_0)$, and $\mathbf{k} = (k \sin \theta, 0, k \cos \theta)$, as shown in Figure 6.6. Working out the cross-products in equation (6.5.12), dividing by k^2 , and factoring out B_0^2 , we obtain the equation

$$\left(\frac{\omega^2}{k}\right) \begin{bmatrix} \tilde{U}_x \\ \tilde{U}_y \\ \tilde{U}_z \end{bmatrix} = V_A^2 \begin{bmatrix} \tilde{U}_x & \tilde{U}_x \\ \tilde{U}_y \cos^2 \theta & 0 \\ 0 & \tilde{U}_x \sin \theta \cos \theta + \tilde{U}_z \cos^2 \theta \end{bmatrix} = V_S^2 \begin{bmatrix} \tilde{U}_x \sin^2 \theta + \tilde{U}_z \sin \theta \cos \theta \\ 0 \\ \tilde{U}_x \sin \theta \cos \theta + \tilde{U}_z \cos^2 \theta \end{bmatrix}, \quad (6.5.13)$$

where the quantity

$$V_A = B_0 / \sqrt{\mu_0 \rho_{m0}} \quad (6.5.14)$$

has units of velocity and is called the Alfvén velocity, after Alfvén [1942]. Substituting $v_p = \omega/k$ for the phase velocity, the homogeneous equation (6.5.13) can be

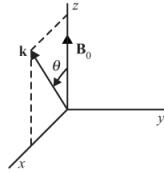


Fig. 6.6 The coordinate system used to analyze MHD wave propagation.

written in matrix form as

$$\begin{bmatrix} v_p^2 - V_S^2 \sin^2 \theta - V_A^2 & 0 & -V_S^2 \sin \theta \cos \theta \\ 0 & v_p^2 - V_A^2 \cos^2 \theta & 0 \\ -V_S^2 \sin \theta \cos \theta & 0 & v_p^2 - V_S^2 \cos^2 \theta \end{bmatrix} \begin{bmatrix} \tilde{U}_x \\ \tilde{U}_y \\ \tilde{U}_z \end{bmatrix} = 0. \quad (6.5.15)$$

This equation has non-trivial solutions for \tilde{U} if and only if the determinant of the matrix is zero, which gives the dispersion relation

$$D(k, \omega) = (v_p^2 - V_A^2 \cos^2 \theta) [v_p^4 - v_p^2 (V_A^2 + V_S^2) + V_A^2 V_S^2 \cos^2 \theta] = 0. \quad (6.5.16)$$

It is easy to show that the dispersion relation has three roots:

$$v_p^2 = \frac{1}{2} (V_A^2 + V_S^2) - \frac{1}{2} [(V_A^2 - V_S^2)^2 + 4V_A^2 V_S^2 \sin^2 \theta]^{1/2}, \quad (6.5.17)$$

$$v_p^2 = V_A^2 \cos^2 \theta, \quad (6.5.18)$$

$$v_p^2 = \frac{1}{2} (V_A^2 + V_S^2) - \frac{1}{2} [(V_A^2 - V_S^2)^2 + 4V_A^2 V_S^2 \sin^2 \theta]^{1/2}. \quad (6.5.19)$$

The roots, equations (6.5.17), (6.5.18) and (6.5.19), are called the slow magnetosonic mode, the transverse Alfvén mode (also called the shear Alfvén mode), and the fast magnetosonic mode, respectively. In the low-temperature low-frequency limit, it can be shown that for $\theta = 0$ the transverse Alfvén mode corresponds to the $n^2 = L$ mode, and the fast

magnetosonic mode corresponds to the $n^2 = R$ mode (see Section ??). The slow magnetosonic mode disappears completely in the limit of zero temperature (i.e., $V_S = 0$). The phase velocities of the three modes are shown as a function of the wave normal angle in Figure 6.7. Two cases, $V_A > V_S$ and $V_A < V_S$, must be considered. For $V_A > V_S$, the transverse

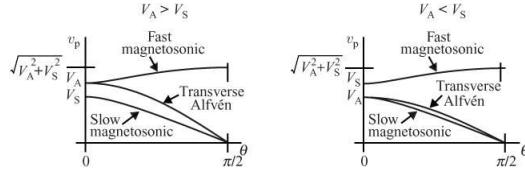


Fig. 6.7 Plots of the phase velocities for the three MHD modes as a function of the wave normal angle for two cases, $V_A > V_S$ and $V_A < V_S$.

Alfvén mode connects with the fast magnetosonic mode at $\theta = 0$, whereas for $V_A < V_S$ the transverse Alfvén mode connects with the slow magnetosonic mode at $\theta = 0$.

Next, we discuss the eigenvectors associated with each of these modes.

6.5.1 The transverse (or shear) Alfvén mode

It is easy to verify that the root for the transverse Alfvén mode, given by equation (6.5.18), has eigenvectors

$$\tilde{\mathbf{U}} = (0, \tilde{U}_y, 0), \quad (6.5.20)$$

$$\tilde{\mathbf{B}} = (0, \tilde{B}_y, 0), \quad \tilde{B}_y = -B_0(\tilde{U}_y/V_A) \operatorname{Sign}(\cos \theta), \quad (6.5.21)$$

$$\tilde{\mathbf{E}} = (\tilde{E}_x, 0, 0), \quad \tilde{E}_x = -B_0\tilde{U}_y, \quad (6.5.22)$$

$$\tilde{\rho}_m = 0. \quad (6.5.23)$$

These eigenvectors are shown in Figure 6.8. As can be seen, the fluid motions for this mode are entirely transverse, with no compressional component (i.e., $\mathbf{k} \cdot \tilde{\mathbf{U}} = 0$). This is the reason the mode is called the transverse Alfvén mode. The propagation velocity is controlled entirely by the Alfvén speed. Since there is no compression, the fluid pressure and temperature play no role in the propagation of this mode. From the direction of the electric and magnetic field, it is easy to see that the Poynting flux, $\mathbf{S} = (1/\mu_0) \mathbf{E} \times \mathbf{B}$, is parallel to the static magnetic field \mathbf{B}_0 . The electromagnetic energy flow is then exactly along the static magnetic field, *independent* of the wave normal angle. From the dispersion relation (6.5.18), it is easy to show that the group velocity is given by

$$\mathbf{v}_g = \nabla_{\mathbf{k}} \omega = V_A \hat{\mathbf{z}}, \quad (6.5.24)$$

which is parallel to the static magnetic field, consistent with the fact that the Poynting flux is parallel to the magnetic field.

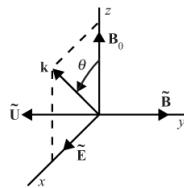


Fig. 6.8 The eigenvectors for the transverse Alfvén mode.

The propagation of the transverse Alfvén wave has an interesting analogy with the propagation of waves on a taut string. From the magnetic pressure tensor we have seen that the magnetic pressure can be thought of as an isotropic pressure plus a tension force per unit area, B^2/μ_0 , along the magnetic field. It is well known that the velocity of propagation of a wave along a taut string is given by

$$v_p = \sqrt{T/\lambda_m}, \quad (6.5.25)$$

where T is the tension and λ_m is the mass per unit length. If we substitute $T = (B^2/\mu_0)\Delta A$ for the tension force and $\lambda_m = \rho_m\lambda A$ for the mass per unit length (where ΔA is the cross-sectional area), the velocity of propagation is given by $v_p = B/\sqrt{\mu_0\rho_m}$, which is just the Alfvén velocity. Thus, the propagation of the transverse (or shear) Alfvén wave can be thought of as being analogous to the propagation of a wave on a taut string with the magnetic field providing the tension and the MHD fluid providing the mass per unit length. This analogy also suggests that waves on different magnetic field lines propagate independently, as though they were on separate strings. Thus, even though the wave vector makes a substantial angle to the magnetic field, as illustrated in Figure 6.9, the wave energy is transported along the magnetic field lines, just as though the wave were propagating on a system of taut strings.

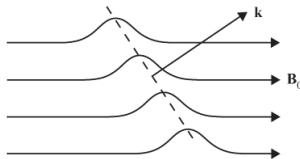


Fig. 6.9 The taut string analogy for the propagation of the transverse Alfvén mode.

6.5.2 The fast and slow magnetosonic modes

As can be seen from the dispersion relations for the slow and fast magnetosonic modes, equations (6.5.17) and (6.5.19) respectively, these modes involve both the magnetic field (via the Alfvén speed) and the plasma pressure (via the sound speed). For this reason these two modes are called the magnetosonic modes. For oblique angles of propagation, it can be shown that the flow velocities have components along the \mathbf{k} vector. Thus, in contrast to the transverse Alfvén mode, which is purely transverse, the magnetosonic modes have both

transverse and longitudinal (i.e., compressional) components. Since the dispersion relation is rather complicated for arbitrary angles of propagation, we first consider the limiting cases of $\theta = 0$ and $\theta = \pi/2$.

For $\theta = 0$ the homogeneous equation is

$$\begin{bmatrix} v_p^2 - V_A^2 & 0 \\ 0 & v_p^2 - V_S^2 \end{bmatrix} \begin{bmatrix} \tilde{U}_x \\ \tilde{U}_z \end{bmatrix} = 0. \quad (6.5.26)$$

The dispersion relation is easily seen to have two roots, $v_p^2 = V_A^2$ and $v_p^2 = V_S^2$. Just which of these two roots corresponds to the fast mode and which corresponds to the slow mode depends on the relative magnitudes of V_A and V_S . For $V_A > V_S$, $v_p^2 = V_A^2$ is the fast mode and $v_p^2 = V_S^2$ is the slow mode. For $V_S > V_A$ the fast and slow designations are reversed. It is easy to show that the eigenvectors corresponding to the $v_p^2 = V_A^2$ root are $\tilde{\mathbf{U}} = (\tilde{U}_x, 0, 0)$, $\tilde{\mathbf{B}} = (\tilde{B}_x, 0, 0)$, $\tilde{\mathbf{E}} = (0, \tilde{E}_y, 0)$ and $\tilde{\rho}_m = 0$. Note that \mathbf{k} and $\tilde{\mathbf{U}}$ are orthogonal. Since $\mathbf{k} \cdot \tilde{\mathbf{U}} = 0$ there is no density compression. Thus, the eigenvectors for this mode have the properties of a *transverse electromagnetic wave*. (That this mode turns out to be transverse is a special case that only occurs at $\theta = 0$.) In contrast, the eigenvectors corresponding to the $v_p^2 = V_S^2$ root are $\tilde{\mathbf{U}} = (0, 0, \tilde{U}_z)$, $\tilde{\mathbf{B}} = (0, 0, 0)$, $\tilde{\mathbf{E}} = (0, 0, 0)$, and $\tilde{\rho}_m = \rho_{m0}(\tilde{U}_z/v_p)$. Note that the electromagnetic fields are zero and that there is a density compression. These eigenvectors have the properties of a *sound wave*. (That the electromagnetic component disappears completely is also a special case that only occurs at $\theta = 0$.)

For $\theta = \pi/2$ the homogeneous equation is

$$\begin{bmatrix} v_p^2 - (V_A^2 + V_S^2) & 0 \\ 0 & v_p^2 \end{bmatrix} \begin{bmatrix} \tilde{U}_x \\ \tilde{U}_z \end{bmatrix} = 0. \quad (6.5.27)$$

In this case, the dispersion relation has only one non-trivial root, $v_p^2 = V_A^2 + V_S^2$. The eigenvectors corresponding to this root are $\tilde{\mathbf{U}} = (\tilde{U}_x, 0, 0)$, $\tilde{\mathbf{B}} = (0, 0, \tilde{B}_z)$, $\tilde{\mathbf{E}} = (0, \tilde{E}_y, 0)$, and $\tilde{\rho}_m = \rho_{m0}(\tilde{U}_x/v_p)$. These eigenvectors share features of both an electromagnetic wave and a sound wave.

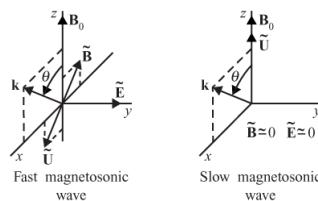


Fig. 6.10 The eigenvectors for the fast and slow magnetosonic modes.

For intermediate wave normal angles, the two magnetosonic modes have both longitudinal (i.e., sound wave) and transverse (i.e., electromagnetic) components. Further insight into the nature of the magnetosonic waves at intermediate wave normal angles can be gained by considering the limit $V_S^2 \ll V_A^2$. Under these conditions it is easy to show that the magnetosonic part of the dispersion relation is approximately

$$D(k, \omega) = (v_p^2 - V_A^2)(v_p^2 - V_S^2 \cos^2 \theta) = 0. \quad (6.5.28)$$

As shown in Figure 6.10, the isotropic root $v_p^2 = V_A^2$ has eigenvectors corresponding to a nearly transverse electromagnetic wave with only a small longitudinal (i.e., sound wave) component, and the anisotropic root $v_p^2 = V_S^2 \cos^2 \theta$ has eigenvectors corresponding to a sound wave with only a small electromagnetic component ($\tilde{\mathbf{E}} \approx 0$ and $\tilde{\mathbf{B}} \approx 0$).

Since the plasma pressure is much smaller than the magnetic field pressure when $V_S^2 \ll V_A^2$, the fluid motion for the slow magnetosonic wave is constrained to be nearly parallel to the static magnetic field. The group velocity and wave energy flow are also nearly parallel to the static magnetic field. The particle motions and wave energy flow for this mode are then similar to the motion of the gas in a parallel array of pipes, with the magnetic field playing the role of the pipes.

6.5.3 MHD wave observations

As mentioned earlier, the existence of MHD waves was first predicted theoretically by Alfvén in 1942. It did not take long before these waves were observed in the laboratory, first in experiments with magnetized conducting fluids, such as mercury [Lundquist, 1949], and later in magnetized plasmas [Wilcox *et al.*, 1960]. It is now widely recognized that MHD waves are of considerable importance, particularly in space plasmas where the ideal MHD condition is most easily satisfied. We will discuss two such examples.

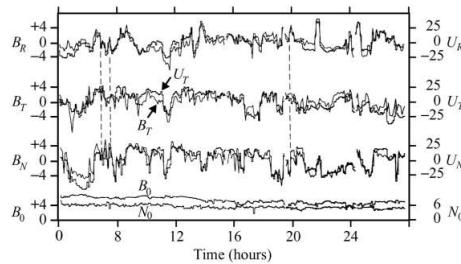


Fig. 6.11 Measurements of the solar wind magnetic field, B , and the plasma flow velocity, U , in an orthogonal (N, R, T) coordinate system [from Davis, 1972]. The fluctuations are due to transverse Alfvén waves propagating outward from the Sun.

The first example occurs in the solar wind. When spacecraft borne magnetic field and plasma measurements became available in the solar wind, it soon became apparent that the solar wind is dominated by large amplitude low-frequency fluctuations. By comparing the magnetic field with the plasma flow velocity, Belcher *et al.* [1969] showed that the fluctuations are caused by Alfvén waves. An example of these fluctuations is shown in Figure 6.11. The top three panels show three components of the solar wind magnetic field, B_N, B_T, B_R , superposed on suitably scaled components of the solar wind velocity, U_N, U_T, U_R . A running average has been subtracted from the measurements in order to reveal the fluctuations. The bottom panel shows the zero-order magnetic field strength,

B_0 , and number density, N_0 . As one can see, the corresponding components of the magnetic field and flow velocity are closely correlated, as would be expected for the transverse Alfvén mode, see equations (6.5.20) and (6.5.21). The almost total absence of comparable fluctuations in the zero-order magnetic field strength and number density uniquely identifies the mode of propagation as the transverse Alfvén mode. The in-phase nature of the correlation also shows that the waves are propagating outward from the Sun. Turbulent convective motions in the photosphere of the Sun are believed to be the source of these waves. Since the electric field of the transverse mode is always perpendicular to the static magnetic field, see Figure 6.8, this mode is almost totally unaffected by the various collisionless damping processes that can exist in a hot, magnetized plasma, see Chapter ???. Thus, once generated, the waves can propagate great distances with little or no attenuation. In contrast, when hot plasma effects are considered, the fast and slow magnetosonic modes always have a small component of the electric field along the static magnetic field (except for the degenerate case at $\theta = 0$), which leads to damping by the thermal plasma.

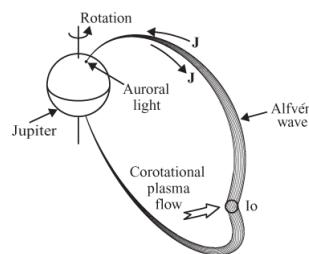


Fig. 6.12 A model showing the excitation of transverse Alfvén waves by Jupiter's moon Io.

The second example is associated with Jupiter's moon Io. In the process of studying radio emissions from Jupiter, Bigg [1964] discovered that Io played a dominant role in controlling radio emissions from Jupiter in the frequency range from about 10 to 40 MHz. To explain the origin of these Io-controlled radio emissions, Goldreich and Lynden-Bell [1969] proposed that Io is a conductor, and that the electromotive force induced by Io's motion through the co-rotating plasma of Jupiter's magnetosphere excites transverse Alfvén waves propagating northward and southward from the moon as shown in Figure 6.12. Because of the finite propagation velocity, the Alfvén wave pattern is swept backward relative to the incident plasma flow, forming a v-shaped standing wave similar to the bow wave of a ship. The electrical currents associated with these Alfvén waves create two quasi-field-aligned current loops, one in each hemisphere, that link Io to the upper atmosphere of Jupiter. The electrons that carry these currents are believed to be responsible for generating the radio emissions. Magnetic field measurements by the Voyager 1 spacecraft during a close flyby of Io have since confirmed the existence of these Alfvén waves [Ness *et al.*, 1979]. Voyager 1 also showed that Io has active volcanism. Gases from these volcanos are now known to be a crucial factor in the Alfvén wave excitation by providing an electrically conducting atmosphere around Io, and also via the momentum they transfer to the plasma when the particles are ionized. The Hubble Space Telescope has also provided dramatic confirmation of the existence of these Alfvén waves by detecting auroral light emission

from the points where the northward and southward propagating Alfvén waves interact with Jupiter's atmosphere [Connerney *et al.*, 1993].

6.6 Static MHD equilibrium

A static equilibrium is defined as a condition in which the fluid is at rest (i.e., $\mathbf{U} = 0$ in some appropriate frame of reference). Since static MHD configurations are frequently encountered in both laboratory and space plasmas, it is useful to examine the conditions required to achieve a static equilibrium. In an ordinary fluid, a static equilibrium is achieved by balancing the gravitational force with an appropriate pressure gradient force. In an MHD fluid, magnetic forces must also be considered. Since magnetic forces are anisotropic, the equilibrium conditions are necessarily considerably more complicated than for an ordinary fluid.

6.6.1 The virial theorem

Before considering the conditions required to achieve a static equilibrium, we first prove an important theorem called the virial theorem. The virial theorem is an exact result that shows that it is impossible for an isolated MHD fluid to generate the magnetic forces necessary to maintain a static equilibrium. The importance of this theorem is obvious. An isolated ball of plasma cannot maintain itself in a static equilibrium just due to internally generated forces. To confine a plasma, it is necessary to provide an external force.

The virial theorem can be proved as follows. We start by writing the momentum equation (6.2.8) in index notation

$$\rho_m \frac{dU_s}{dt} = - \sum_r \frac{\partial}{\partial x_r} (P_{rs} + T_{rs}) \quad (6.6.1)$$

If a static equilibrium exists ($U_s = 0$), the left-hand side of the above equation must be zero. Next, we multiply the above equation by x_s , sum over all three coordinates, and integrate over all space to obtain

$$0 = - \sum_{r,s} \int_V x_s \frac{\partial}{\partial x_r} (P_{rs} + T_{rs}) d^3x. \quad (6.6.2)$$

The terms on the right-hand side of this equation can be integrated by parts once to give

$$0 = - \sum_{r,s} \int_S x_s (P_{rs} + T_{rs}) dA_r + \sum_{r,s} \int_V \frac{\partial x_s}{\partial x_r} (P_{rs} + T_{rs}) d^3x. \quad (6.6.3)$$

Note that the first term involves a surface integral at infinity. For an isolated three-dimensional system, we can assume that the surface integral is zero, the only requirement being

that P_{rs} and T_{rs} decrease sufficiently rapidly at infinity (i.e., faster than $1/R^3$). Using the identity $\partial x_s/\partial x_r = \delta_{rs}$, where δ_{rs} is the Kronecker delta, equation (6.6.3) simplifies to

$$0 = \int_V \text{Trace}(\overset{\leftrightarrow}{\mathbf{P}} + \overset{\leftrightarrow}{\mathbf{T}}) d^3x, \quad (6.6.4)$$

where the trace is the sum of the diagonal terms in the tensor. However, it is easy to see that the traces of $\overset{\leftrightarrow}{\mathbf{P}}$ and $\overset{\leftrightarrow}{\mathbf{T}}$ are positive definite,

$$\begin{aligned} \text{Trace}\overset{\leftrightarrow}{\mathbf{P}} &= \sum_{r,s} m_s \int_V (v_r - U_r)^2 f_s(\mathbf{v}) d^3v \\ \text{Trace}\overset{\leftrightarrow}{\mathbf{T}} &= B^2/2\mu_0, \end{aligned} \quad (6.6.5)$$

so the right-hand side of equation (6.6.4) cannot be zero. Therefore, the original assumption, that the system is in static equilibrium, must be false. A static equilibrium for an isolated MHD system is impossible.

It is important to note the conditions for which the virial theorem applies. The plasma must be finite in three dimensions and must be completely isolated, with no external forces. The theorem does not apply to systems that are infinite in one or two dimensions, or to systems that contain rigid current-carrying conductors (due to the external mechanical forces on the conductors).

6.6.2 Conditions for a static MHD equilibrium

If we choose a reference frame in which the fluid is at rest, the conditions for a static equilibrium are obtained by simply setting $\partial/\partial t = 0$ and $\mathbf{U} = 0$ in the MHD equations given in Section 6.1.5. The relevant equations are then

$$\mathbf{J} \times \mathbf{B} = \nabla P, \quad (6.6.6)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}, \quad (6.6.7)$$

and

$$\nabla \cdot \mathbf{B} = 0, \quad (6.6.8)$$

where the plasma pressure, P , is assumed to be isotropic. By taking the dot-product of equation (6.6.6) with \mathbf{J} and \mathbf{B} , it follows that $\mathbf{J} \cdot \nabla P = 0$ and $\mathbf{B} \cdot \nabla P = 0$. Since $\nabla P = 0$ is perpendicular to a surface of constant pressure, it also follows that the current density \mathbf{J} and the magnetic field \mathbf{B} must lie on surfaces of constant pressure.

Static equilibria can be divided into two classes, depending upon whether the $\mathbf{J} \times \mathbf{B}$ force vanishes or not. If $\mathbf{J} \times \mathbf{B} = 0$, the equilibrium is said to be force free. If $\mathbf{J} \times \mathbf{B} \neq 0$, the equilibrium is said to be force balanced.

6.6.3 Force-free equilibria

Force-free MHD equilibria typically occur in very tenuous plasmas where the pressure gradient forces are so small that they can be neglected. Since the condition for a force-free equilibrium is $\mathbf{J} \times \mathbf{B} = 0$, it is clear that the force-free condition is satisfied whenever $\mu_0 \mathbf{J} = \nabla \times \mathbf{B} = \alpha \mathbf{B}$, where α is a scalar function.

To understand the implications of the equation $\nabla \times \mathbf{B} = \alpha \mathbf{B}$, we must investigate the function α . We start by noting that the divergence of the curl of any vector field is zero, so that

$$\nabla \cdot (\nabla \times \mathbf{B}) = \alpha(\nabla \cdot \mathbf{B}) + \mathbf{B} \cdot \nabla \alpha = 0. \quad (6.6.9)$$

Since $\nabla \cdot \mathbf{B} = 0$, this equation simplifies to

$$\mathbf{B} \cdot \nabla \alpha = 0. \quad (6.6.10)$$

Since $\nabla \alpha$ is perpendicular to a surface of constant α , the above equation shows that the magnetic field lines must lie on a constant- α surface. Next we show that the constant- α surface cannot be a simple closed surface. To prove this relationship, let C be a closed curve everywhere parallel to the magnetic field. The line integral of \mathbf{B} along this closed path clearly cannot be zero,

$$\int_C \mathbf{B} \cdot d\ell \neq 0. \quad (6.6.11)$$

Using Stokes' theorem and $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} = \alpha \mathbf{B}$, the above integral can be written

$$\int_C \mathbf{B} \cdot d\ell = \int_S (\nabla \times \mathbf{B}) \cdot d\mathbf{A} = \int_S \alpha \mathbf{B} \cdot d\mathbf{A} \neq 0, \quad (6.6.12)$$

where S is a simple surface bounded by the curve C . Next, distort the surface S such that it coincides with the constant- α surface. This procedure can only be performed if the surface is simply closed, like a sphere. Otherwise, Stokes' theorem is not valid. Since the function α is constant on this surface, α can be taken outside of the integral, so that

$$\int_C \mathbf{B} \cdot d\ell = \alpha \int_S \mathbf{B} \cdot d\mathbf{A} \neq 0. \quad (6.6.13)$$

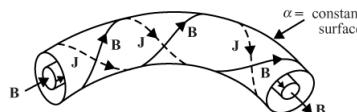


Fig. 6.13

The magnetic field and current for a force-free equilibrium form helix-like lines around torus shaped surfaces of constant α .

However, earlier we showed that the magnetic field lines must lie on the constant- α surface,

so $\mathbf{B} \cdot d\mathbf{A} = 0$. Thus, the integral on the right-hand side of the above equation is zero, which leads to a contradiction. The constant- α surface cannot be a simply closed surface.

Using a theorem from topology known as Hopf's theorem [Alexandroff and Hopf, 1935], it can be shown that the simplest topological form for a constant- α surface that satisfies the above restrictions is a torus. Hopf's theorem states "if a non-vanishing, bounded vector field (\mathbf{J} and \mathbf{B} , in our case) lies on a smooth surface (the constant- α surface), then the surface must be toroidal." The current and magnetic field must then form helix-like lines around the torus as shown in Figure 6.13. This result is quite general, and also applies to force balanced MHD equilibria, since equation (6.6.6) implies that $\mathbf{B} \cdot \nabla P = 0$. The surfaces of constant P then play the same role as surfaces of constant α do for force-free equilibria.

Particularly interesting solutions to the static equilibrium equations can be found for two-dimensional geometries with cylindrical symmetry. For such a geometry, it is easy to show that the equation $\nabla \times \mathbf{B} = \alpha \mathbf{B}$ is satisfied by a magnetic field $\mathbf{B} = [0, B_\phi, B_z]$ of the form

$$B_\phi = \frac{B_0 k \rho}{1 + k^2 \rho^2}, \quad (6.6.14)$$

and

$$B_z = \frac{B_0}{1 + k^2 \rho^2}. \quad (6.6.15)$$

In this case, the z component of the current density is given by

$$J_z = \frac{2k B_0 / \mu_0}{(1 + k^2 \rho^2)^2}, \quad (6.6.16)$$

and the function α is given by

$$\alpha = \frac{\mu_0 J_z}{B_z} = \frac{2k}{1 + k^2 \rho^2}. \quad (6.6.17)$$



Fig. 6.14 A flux rope consisting of helical magnetic field lines nested around a central axis.

The magnetic field lines consist of helices that spiral around the central axis with pitch angles that vary from 0 along the central axis to $\pi/2$ at large distances from the central axis. Such a magnetic field configuration is often called a flux rope (see Figure 6.14).

Magnetic field geometries of the type described above have been observed in the ionosphere of Venus, in the Earth's magnetosphere, and in the solar corona. During disturbed periods large currents often flow along magnetic field lines from one region of the Sun to another. These currents often heat the plasma to such a high temperature that the magnetic field lines can be detected optically from the Earth. The field lines typically consist of an arch with a helical structure, such as shown in Figure 6.15.

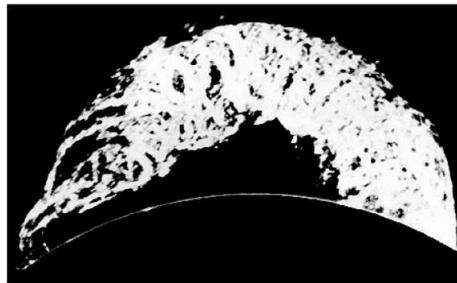


Fig. 6.15 The Sun's magnetic field often has helical flux rope-like structures [courtesy of High Altitude Observatory].

Another class of force-free solutions with cylindrical symmetry is of the form

$$B_\phi = B_0 J_1(\alpha\rho), \quad (6.6.18)$$

and

$$B_z = B_0 J_0(\alpha\rho), \quad (6.6.19)$$

where α and B_0 are positive constants and J_0 and J_1 are Bessel functions of order 0 and 1. Note that for a cylindrical plasma of radius a , equation (6.6.19) leads to a reversal of the axial field at the first zero of $J_0(x)$, i.e., at $\alpha\rho = 2.404, 5.520$, etc. Configurations of this type are known as reverse field pinches (RFPs), and are among the leading candidates for magnetic confinement of fusion plasmas.

Taylor [1974] proposed that reverse field pinches are generated spontaneously when a turbulent plasma relaxes to a state of minimum potential energy subject to the constancy of a quantity called the magnetic helicity, which is defined by

$$\mathcal{H} = \int_V \mathbf{A} \cdot \mathbf{B} d^3x, \quad (6.6.20)$$

where \mathbf{A} is the vector potential, i.e., $\mathbf{B} = \nabla \times \mathbf{A}$. Magnetic helicity is an exact invariant for a plasma that obeys the frozen field condition (6.3.7) or equivalently, $\mathbf{E} + \mathbf{U} \times \mathbf{B} = 0$. The invariance of \mathcal{H} can be shown by considering the time derivative of equation (6.6.20), which is given by

$$\frac{d\mathcal{H}}{dt} = \int_V \left(\frac{\partial \mathbf{A}}{\partial t} \cdot \mathbf{B} + \mathbf{A} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) d^3x. \quad (6.6.21)$$

Using Faraday's law (6.1.30), the electric field, \mathbf{E} , can be written in terms of a scalar potential, ϕ , and the vector potential, \mathbf{A} , as

$$\mathbf{E} = -\nabla\phi - \frac{\partial \mathbf{A}}{\partial t}. \quad (6.6.22)$$

Substituting equation (6.6.22) for $\partial\mathbf{A}/\partial t$ and equation (6.1.30) for $\partial\mathbf{B}/\partial t$ into equation (6.6.21), we obtain

$$\frac{d\mathcal{H}}{dt} = \int_V \mathbf{E} \cdot \mathbf{B} d^3x - \int_V \mathbf{B} \cdot \nabla\Phi d^3x - \int_V (\nabla \times \mathbf{E}) \cdot \mathbf{A} d^3x. \quad (6.6.23)$$

Using the identities

$$\nabla \cdot (\Phi\mathbf{B}) = \mathbf{B} \cdot \nabla\Phi + \Phi(\nabla \cdot \mathbf{B}) = \mathbf{B} \cdot \nabla\Phi, \quad (6.6.24)$$

and

$$\begin{aligned} \nabla \cdot (\mathbf{E} \times \mathbf{A}) &= (\nabla \times \mathbf{E}) \cdot \mathbf{A} - \mathbf{E} \cdot \nabla \times \mathbf{A} \\ &= (\nabla \times \mathbf{E}) \cdot \mathbf{A} - \mathbf{E} \cdot \mathbf{B}, \end{aligned} \quad (6.6.25)$$

we can rewrite equation (6.6.23) in the form

$$\frac{d\mathcal{H}}{dt} = -2 \int_V \mathbf{E} \cdot \mathbf{B} d^3x - \int_V \nabla \cdot (\Phi\mathbf{B}) d^3x - \int_V \nabla \cdot (\mathbf{E} \times \mathbf{A}) d^3x. \quad (6.6.26)$$

The second and third volume integrals on the right-hand side of the above equation can be converted to surface integrals using Gauss' theorem. This gives

$$\frac{d\mathcal{H}}{dt} = -2 \int_V \mathbf{E} \cdot \mathbf{B} d^3x - \int_S \hat{\mathbf{n}} \cdot (\mathbf{B}\Phi + \mathbf{E} \times \mathbf{A}) dA, \quad (6.6.27)$$

where $\hat{\mathbf{n}}$ is the unit normal (pointing outward) at the area element dA on the surface S . Assuming that the plasma is bounded by a perfectly conducting wall that is identified with the surface S , the \mathbf{E} and \mathbf{B} fields obey the boundary conditions

$$\hat{\mathbf{n}} \times \mathbf{E} = 0, \quad \hat{\mathbf{n}} \cdot \mathbf{B} = 0. \quad (6.6.28)$$

From these boundary conditions, it is easy to see that the surface integral on the right-hand side of equation (6.6.27) vanishes. Also, because in an ideal plasma ($\mathbf{E} + \mathbf{U} \times \mathbf{B} = 0$) we have $\mathbf{E} \cdot \mathbf{B} = 0$, it follows that the volume integral on the right-hand side of equation (6.6.27) also vanishes, so $d\mathcal{H}/dt = 0$. Thus, we have shown that the magnetic helicity is a global invariant, i.e.,

$$\mathcal{H} = \text{constant}. \quad (6.6.29)$$

The proof of the invariance of \mathcal{H} given above relies on the assumption that the plasma is ideal, which no plasmas are, in reality. Taylor [1974] conjectured that in a turbulent non-ideal plasma the helicity remains an approximate constant while the potential energy, given by equation (6.4.18), decays to a minimum value. It can be shown that these minimum energy equilibria are force free and obey the equations $\nabla \times \mathbf{B} = \alpha \mathbf{B}$ and $P = \text{constant}$, where α is a constant. In a cylindrical geometry, assuming that equilibrium quantities depend only on ρ , one then obtains the Bessel function solutions (6.6.18) and (6.6.19).

Taylor's theory has been successful not only in accounting for the dominant behavior of reverse field pinches, but also has found applications in space and astrophysical plasmas. For further discussions of this interesting subject, see Taylor [1986] and Priest and Forbes [2000].

6.6.4 Force balanced equilibria

If the plasma is not sufficiently tenuous for the force-free condition to apply, then the $\mathbf{J} \times \mathbf{B}$ force must be balanced by the pressure gradient force. Using the momentum equation (6.2.10) with $\mathbf{U} = 0$, the force balance condition can be written in the form

$$\nabla \left(P + \frac{B^2}{2\mu_0} \right) = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B}, \quad (6.6.30)$$

where the gradient terms are collected on the left-hand side and the non-gradient terms are on the right-hand side. Since there is an infinite number of solutions to this equation, to proceed further we must limit the discussion to certain magnetic field geometries that are of special interest.

6.6.4.1 Cylindrical geometries

We start by discussing two-dimensional cylindrical geometries that are characterized by a magnetic field of the form $\mathbf{B} = [0, B_\phi(\rho), B_z(\rho)]$. For such a magnetic field, it is easy to show that the radial component of equation (6.6.30) becomes

$$\frac{d}{d\rho} \left(P + \frac{B_\phi^2}{2\mu_0} + \frac{B_z^2}{2\mu_0} \right) = -\frac{B_\phi^2}{\mu_0 \rho}. \quad (6.6.31)$$

Notice that the radial component of the $(\mathbf{B} \cdot \nabla) \mathbf{B}$ force is entirely controlled by the ϕ component of the magnetic field. This force is caused by the magnetic tension. Because of the curvature of the magnetic field the magnetic tension force is directed radially inward. If the magnetic field is known, the components of the current density can be calculated from Ampère's law (6.1.28), which, in cylindrical coordinates, is given by

$$J_\rho = 0, \quad (6.6.32)$$

$$J_\phi = -\frac{1}{\mu_0} \frac{dB_z}{d\rho}, \quad (6.6.33)$$

$$J_z = \frac{1}{\mu_0 \rho} \frac{d}{d\rho} (\rho B_\phi). \quad (6.6.34)$$

Equation (6.6.31) involves three functions, $B_\phi(\rho)$, $B_z(\rho)$, and $P(\rho)$. Any two of these functions can be specified, whereupon the third can be determined by solving the resulting differential equation, subject to the appropriate boundary conditions. Since considerable

freedom exists in our choice of the two “free” functions, equation (6.6.31) itself permits an infinite number of possible solutions. In the following, we discuss a few of this very large class of possible solutions.

The Z-pinch In the Z-pinch, also sometimes called a Bennett pinch, a cylindrically symmetric plasma is confined by a purely azimuthal magnetic field $B_\phi(\rho)$. A sketch of the basic field geometry of a Z-pinch is shown in Figure 6.16. The confining magnetic field B_ϕ is entirely generated by the axial current density J_z according to equation (6.6.34). Since $B_z = 0$, equation (6.6.31) reduces to

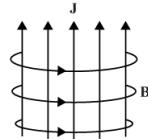


Fig. 6.16 The magnetic field of a Z-pinch is produced internally by a current along the axis of a plasma column.

$$\frac{d}{d\rho} \left(P + \frac{B_\phi^2}{2\mu_0} \right) = -\frac{B_\phi^2}{\mu_0 \rho}. \quad (6.6.35)$$

The above equation permits an infinite number of solutions. Consider a simple case in which the axial current density, J_{z0} , is constant within a plasma column of radius $\rho = a$, and zero outside,

$$\begin{aligned} J_z(\rho) &= J_{z0}, & \text{for } \rho \leq a \text{ and} \\ J_z(\rho) &= 0, & \text{for } \rho > a. \end{aligned} \quad (6.6.36)$$

With $J_z(\rho)$ specified by the above equation, equation (6.6.34) can be solved for $B_\phi(\rho)$ to yield

$$\begin{aligned} B_\phi(\rho) &= \frac{\mu_0 I \rho}{2\pi a^2}, & \text{for } \rho \leq a \text{ and} \\ B_\phi(\rho) &= \frac{\mu_0 I}{2\pi \rho}, & \text{for } \rho > a. \end{aligned} \quad (6.6.37)$$

The total current is given by

$$I = \int_0^a J_z 2\pi \rho d\rho = \pi a^2 J_{z0}. \quad (6.6.38)$$

Substituting expression (6.6.37)) for $B_\phi(\rho)$ into equation (6.6.35) and assuming that the plasma pressure at the edge of the column vanishes, i.e., $P(a) = 0$, it is easy to show that

$$P(\rho) = \mu_0 \left(\frac{1}{2\pi a} \right)^2 \left(1 - \frac{\rho^2}{a^2} \right). \quad (6.6.39)$$

If there are sufficient collisions to maintain an isotropic pressure distribution, one can assume that the pressure is given by the ideal gas law, $P = n\kappa T$. To proceed further, it is convenient to define a line number density given by

$$N_\ell = \int_0^a n 2\pi \rho \, d\rho. \quad (6.6.40)$$

Assuming that the temperature is constant across the plasma column, the above equation can be rewritten as

$$N_\ell \kappa T = \int_0^a n \kappa T 2\pi \rho \, d\rho = 2\pi \int_0^a P \rho \, d\rho. \quad (6.6.41)$$

Integrating the above equation by parts once gives

$$N_\ell \kappa T = [\pi P \rho^2]_0^a - \pi \int_0^a \rho^2 (dP/d\rho) \, d\rho. \quad (6.6.42)$$

If the pressure is assumed to go to zero at the outer boundary of the plasma column, the first term on the right vanishes. Substituting $dP/d\rho = -J_z B_\phi$ into the second term then gives

$$N_\ell \kappa T = \pi \int_0^a \rho^2 J_z B_\phi \, d\rho. \quad (6.6.43)$$

Equation (6.6.34) can then be used to eliminate J_z , which gives

$$N_\ell \kappa T = \frac{\pi}{\mu_0} \int_0^a \rho B_\phi \frac{d}{d\rho}(\rho B_\phi) \, d\rho = \frac{\pi}{2\mu_0} \int_0^a \frac{d}{d\rho}(\rho B_\phi)^2 \, d\rho = \frac{\pi}{2\mu_0} a^2 B_\phi^2(a). \quad (6.6.44)$$

To interpret the above equation, it is useful to note that the total current in the column can be evaluated in terms of the magnetic field, $B_\phi(a)$, at the edge of the column,

$$I = \int_0^a J_z 2\pi \rho \, d\rho = \frac{1}{\mu_0} \int_0^a \frac{d}{d\rho}(\rho B_\phi) 2\pi \rho \, d\rho = \frac{2\pi}{\mu_0} a B_\phi(a). \quad (6.6.45)$$

Combining the above expression with equation (6.6.44) then gives

$$I^2 = \frac{8\pi N_\ell \kappa T}{\mu_0}. \quad (6.6.46)$$

This equation is called the Bennett relation, after Bennett [1934]. Note that the total current is completely independent of the details of the plasma pressure profile.

The Z-pinch can be set up in the laboratory by simply applying a voltage across electrodes at the ends of a plasma column. As will be discussed later in Section 6.7.2, the Z-pinch is unstable and often evolves into periodic sausage-like structures.

The ϕ -pinch The ϕ -pinch is complementary to the Z-pinch in the sense that the only non-vanishing component of the current density is in the ϕ direction, $J_\phi = J_\phi(\rho)$, rather than in the z direction. This current produces a purely axial magnetic field, $B_z = B_z(\rho)$. In practice

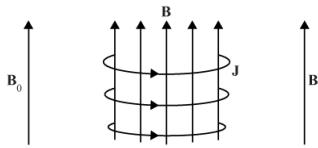


Fig. 6.17 The magnetic field of a ϕ -pinch is externally imposed.

this magnetic field must be imposed by an external boundary condition, usually a constant field, B_0 , parallel to the z axis, as shown in Figure 6.17. For this magnetic field geometry, equation (6.6.31) reduces to

$$\frac{d}{d\rho} \left(P + \frac{B_z^2}{2\mu_0} \right) = 0, \quad (6.6.47)$$

which can be integrated directly to give

$$P + \frac{B_z^2}{2\mu_0} = \text{constant}. \quad (6.6.48)$$

The constant on the right-hand side of the above equation can be determined by assuming that the plasma pressure vanishes at the edge of the column, with a constant externally applied axial field B_0 in the vacuum surrounding the plasma. The above equation then becomes

$$P + \frac{B_z^2}{2\mu_0} = \frac{B_0^2}{2\mu_0}. \quad (6.6.49)$$

This equation also applies even when the plasma column does not have cylindrical symmetry. The only requirement is that the magnetic field lines be straight. The $(\mathbf{B} \cdot \nabla)\mathbf{B}$ term on the right of equation (6.6.30) then vanishes. A self-consistent model illustrating this type of equilibrium was previously discussed in Section ??.

The screw pinch In the screw pinch, both the axial and azimuthal components of the magnetic field are non-zero, that is, $B_\phi \neq 0$ and $B_z \neq 0$. Clearly, the Z-pinch and ϕ -pinch are special cases of the screw pinch. In a screw pinch, the magnetic field lines do not close on themselves, but instead form helical lines around the axis of the cylinder. This is not hard to visualize. Imagine adding a B_z field to the closed B_ϕ lines of a Z-pinch. The result is a set of nested cylindrical surfaces that are covered by field lines that wrap around the cylinders in helical trajectories, like the flux rope shown in Figure 6.14. The magnetic field lines are also sheared in the radial direction. The amount of the shear from one radial position to another depends on the radial dependences of B_ϕ and B_z . We assume that the screw pinch is periodic along z , with a periodicity length $2\pi R$, where R is a constant, and a is the radius of the plasma column. The amount of shear is then measured by the radial dependence of the so-called q -factor, defined by the equation

$$q(\rho) = \frac{\rho B_z}{R B_\phi}. \quad (6.6.50)$$

The q -factor is also sometimes used for the description of a toroidal plasma (see the next section) in which case R and a are the major and minor radii of the plasma torus. In a screw pinch, there are two free functions that need to be specified before a unique equilibrium can be obtained. If we choose $B_\phi(\rho)$ and $B_z(\rho)$ as these two functions, equation (6.6.31) can be integrated to determine the plasma pressure $P(\rho)$. The current densities $J_\phi(\rho)$ and $J_z(\rho)$ can then be obtained from equations (6.6.33) and (6.6.34), respectively.

A simple example of a screw pinch equilibrium is one in which a constant B_z field,

$$B_z = B_0 = \text{constant}, \quad (6.6.51)$$

is added to a Z-pinch. Note that the addition of a constant B_z field makes no contribution to the equilibrium equation (6.6.31). If the B_ϕ field is specified by equation (6.6.37), one obtains the q -profile,

$$\begin{aligned} q(\rho) &= \frac{2\pi B_0 a^2}{\mu_0 I R} && \text{for } \rho \leq a, \quad \text{and} \\ q(\rho) &= \frac{2\pi B_0 \rho^2}{\mu_0 I R} && \text{for } \rho > a. \end{aligned} \quad (6.6.52)$$

Within the plasma column $\rho \leq a$, the q -profile is flat, that is, $dq/d\rho = 0$. The pressure profile for such an equilibrium is given by equation (6.6.39).

6.6.4.2 Toroidal geometries

In Section 6.6.3 we pointed out from Hopf's theorem that the magnetic field lines for a magnetostatic equilibrium must lie on the surface of a torus. We now discuss such force balanced equilibria for an axially symmetric toroidal geometry. In Figure 6.18, we show a cross-section of a plasma torus in which all of the equilibrium quantities depend only on the coordinates ρ and z and are independent of the azimuthal angle ϕ . Such equilibria are of great interest for the tokamak, which is the most widely investigated magnetic confinement device for achieving controlled fusion.

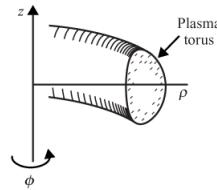


Fig. 6.18 The coordinate system used to analyze a torus geometry.

To analyze toroidal geometries, we must develop some techniques for representing toroidal magnetic fields. The assumption of axial symmetry ($\partial/\partial\phi = 0$) implies that $\nabla \cdot \mathbf{B} = 0$ can be written

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho B_\rho) + \frac{\partial B_z}{\partial z} = 0, \quad (6.6.53)$$

where $\mathbf{B} = [B_\rho, B_\phi, B_z]$. If we represent the magnetic field via a vector potential \mathbf{A} , such that $\mathbf{B} = \nabla \times \mathbf{A}$, it follows that $\nabla \cdot \mathbf{B} = 0$ is automatically satisfied, with B_ρ and B_z components given by

$$B_\rho = -\frac{\partial A_\phi}{\partial z}, \quad (6.6.54)$$

and

$$B_z = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho A_\phi). \quad (6.6.55)$$

The form of these equations motivates the definition of a flux function $\psi(\rho, z) = \rho A_\phi$ such that

$$B_\rho = -\frac{1}{\rho} \frac{\partial \psi}{\partial z}, \quad (6.6.56)$$

and

$$B_z = \frac{1}{\rho} \frac{\partial \psi}{\partial \rho}. \quad (6.6.57)$$

Note that with B_ρ and B_z defined by the above two equations, equation (6.6.53) is satisfied automatically. Furthermore, it follows that

$$\mathbf{B} \cdot \nabla \psi = B_\rho \frac{\partial \psi}{\partial \rho} + B_z \frac{\partial \psi}{\partial z} = 0, \quad (6.6.58)$$

which means that magnetic field lines lie on toroidal surfaces of constant ψ . From the condition $\mathbf{B} \cdot \nabla P = 0$ it also follows that if $\nabla P \neq 0$, then

$$P = P(\psi). \quad (6.6.59)$$

Next, we consider the ϕ component of the force balance equation (6.6.6), which is

$$J_z B_\rho - J_\rho B_z = \frac{1}{\rho} \frac{\partial P}{\partial \phi} = 0. \quad (6.6.60)$$

Using Ampère's law (6.6.7) to express J_z and J_ρ in terms of B_ρ and B_ϕ , the above equation can be rewritten

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho B_\phi) B_\rho + \frac{\partial B_\phi}{\partial z} B_z = 0. \quad (6.6.61)$$

By defining a new function

$$F = \rho B_\phi, \quad (6.6.62)$$

equation (6.6.61) can be written in the simple form

$$\frac{\partial F}{\partial \rho} B_\rho + \frac{\partial F}{\partial z} B_z = 0, \quad (6.6.63)$$

or

$$\mathbf{B} \cdot \nabla F = 0. \quad (6.6.64)$$

Since both $\mathbf{B} \cdot \nabla \psi = 0$ and $\mathbf{B} \cdot \nabla F = 0$, it follows that F can be expressed as a function of ψ

$$F = F(\psi). \quad (6.6.65)$$

Thus, by using two functions, ψ and F , and equations (6.6.56), (6.6.57), and (6.6.62), the magnetic field can be represented in the form

$$\mathbf{B} = \left(\frac{-1}{\rho} \frac{\partial \psi}{\partial z} \right) \hat{\rho} + \left(\frac{1}{\rho} \frac{\partial \psi}{\partial \rho} \right) \hat{\mathbf{z}} + \frac{F(\psi)}{\rho} \hat{\Phi}, \quad (6.6.66)$$

or

$$\mathbf{B} = \left(\frac{\nabla \psi}{\rho} \right) \times \hat{\Phi} + \frac{F(\psi)}{\rho} \hat{\Phi}, \quad (6.6.67)$$

where $\hat{\rho}$, $\hat{\Phi}$, $\hat{\mathbf{z}}$ are the usual unit vectors in the cylindrical coordinates.

Finally, we consider the ρ component of the force balance equation, (6.6.6), which is given by

$$J_\phi B_z - J_z B_\phi = \frac{\partial P}{\partial \rho}. \quad (6.6.68)$$

Following an approach similar to that used for the ϕ component, we proceed by eliminating J_ϕ and J_z using Ampère's law (6.6.7). From the B_ϕ component of Ampère's law, we can obtain an equation for J_ϕ

$$\begin{aligned} \mu_0 J_\phi &= \frac{\partial B_\rho}{\partial z} - \frac{\partial B_z}{\partial \rho} \\ &= -\frac{1}{\rho} \frac{\partial^2 \psi}{\partial z^2} - \frac{1}{\rho} \frac{\partial^2 \psi}{\partial \rho^2} + \frac{1}{\rho^2} \frac{\partial \psi}{\partial \rho} \\ &= -\frac{1}{\rho} \Delta^* \psi, \end{aligned} \quad (6.6.69)$$

where Δ^* is the operator defined by

$$\Delta^* \psi = \rho \frac{\partial}{\partial \rho} \left(\frac{1}{\rho} \frac{\partial \psi}{\partial \rho} \right) + \frac{\partial^2 \psi}{\partial z^2} = \rho^2 \nabla \cdot \left(\frac{\nabla \psi}{\rho^2} \right). \quad (6.6.70)$$

From the z component of Ampère's law (6.6.7), we can obtain an equation for J_z

$$\mu_0 J_z = \frac{1}{\rho} \frac{\partial}{\partial \rho} (\rho B_\phi) = \frac{1}{\rho} \frac{\partial F}{\partial \rho} = \frac{1}{\rho} \frac{dF}{d\phi} \frac{\partial \psi}{\partial \rho}. \quad (6.6.71)$$

By substituting $B_\phi = F(\psi)/\rho$, $\partial P/\partial \rho = (dP/d\psi)(\partial \psi/\partial \rho)$, equation (6.6.57) for B_z , equation (6.6.69) for J_ϕ , and equation (6.6.71) for J_z into equation (6.6.68), one obtains the equation

$$\Delta^* \psi = -\mu_0 \rho^2 \frac{dP}{d\psi} - \frac{1}{2} \frac{d}{d\psi} F^2. \quad (6.6.72)$$

This equation is known as the Grad-Shafranov equation and is widely used to calculate magnetostatic equilibria in axisymmetric toroidal systems [Shafranov, 1966]. The usual strategy to solve this equation is as follows. First specify two free functions, the plasma pressure $P = P(\psi)$ and the toroidal field function $F = F(\psi)$, and solve equation (6.6.72) with specified boundary conditions to determine the flux function $\psi = \psi(\rho, z)$. Once $\psi(\rho, z)$ is known, the three components of the magnetic field can be calculated using equation (6.6.66). The pressure profile can then be obtained from the relation $P = P(\psi(\rho, z))$.

In most circumstances of physical interest, the Grad-Shafranov equation is solved numerically. However, the procedure can be illustrated by considering some exact analytical solutions, known as the Solov'ev equilibria [Solov'ev, 1968].

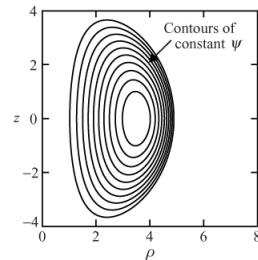


Fig. 6.19 A plot of surfaces of constant ψ for the parameters $C_1 = 1$, $C_2 = -8$, $C_3 = -20$, $C_4 = 20$, and $C_5 = 0.2$.

Assume that the two free functions, $F(\psi)$ and $P(\psi)$, are given by

$$\mu_0 P'(\psi) = -C_2, \quad (6.6.73)$$

and

$$F F'(\psi) = C_1, \quad (6.6.74)$$

where C_1 and C_2 are constants [Freidberg, 1987]. By direct substitution it can be shown that an exact solution of the Grad-Shafranov equation (6.6.72) is given by

$$\psi(\rho, z) = -\frac{C_1}{2}z^2 + \frac{C_2}{8}\rho^4 + C_3 + C_4\rho^2 + C_5(\rho^4 - 4\rho^2z^2), \quad (6.6.75)$$

where C_1 through C_5 are constants. Since the above equation has five arbitrary constants, the equation clearly represents a large family of solutions. Some of these solutions have surfaces of constant ψ that are nested, as shown in Figure 6.19. Solutions of this type are often used in the analysis of toroidal fusion confinement machines, such as the tokamak. Note that, unlike the cylindrical equilibria discussed in Section 6.6.4.1 that depend on only one coordinate (ρ), this equilibrium depends on two coordinates (ρ and z).

6.7 MHD stability

In the previous section we showed that a large variety of solutions to the static MHD equilibrium equations can be found. However, the existence of a solution to the equilibrium equations does not prove that the configuration can actually exist in nature, since the equilibrium may be unstable. If the equilibrium is unstable then the system will start to evolve in time, and will either evolve to some other completely different equilibrium configuration, or never reach a steady state. In order to understand which static MHD equilibria can actually exist in nature it is important to consider the issue of stability. In this section we discuss some of the techniques used to analyze the stability of MHD equilibria.

To provide an introduction to the techniques involved, we start by discussing some simple configurations that can be shown to be unstable using somewhat restrictive assumptions. We then proceed to develop two methods, called (i) the normal mode method, and (ii) the energy method, that can be used to test whether an arbitrary MHD equilibrium is stable or unstable.

6.7.1 MHD stability: a mechanical analogy

It is useful to start the discussion of stability by considering a simple mechanical analog, which is the motion of a particle in a one-dimensional potential energy function, $W(x)$, such as the one shown in Figure 6.20. For such a system, static equilibrium solutions exist at all points for which the force $F_x = -\partial W/\partial x = 0$. To test the stability at a given equilibrium point the procedure is to give the particle a small displacement, or perturbation. If the particle returns to the equilibrium point the system is stable. If it does not return it is unstable. From simple physical considerations one can see that the condition for stability is that the equilibrium point must be at a local minimum in the potential energy. To prove this assertion formally, expand the potential energy in a Taylor series around one of the equilibrium points. Since $\partial W/\partial x = 0$ at the equilibrium point, then to lowest order in the displacement, Δx , the change in the potential energy relative to the potential energy at the equilibrium point is given by

$$\delta W = \frac{1}{2} \left(\frac{\partial^2 W}{\partial x^2} \right) \Delta x^2, \quad (6.7.1)$$

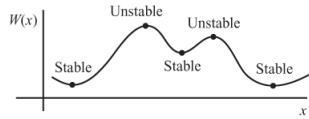


Fig. 6.20 | The potential energy of a mechanical system has points of stable equilibrium, characterized by $\partial^2 W / \partial x^2 > 0$, and points of unstable equilibrium, characterized by $\partial^2 W / \partial x^2 < 0$.

where $(\partial^2 W / \partial x^2)$ is evaluated at the equilibrium point. For such a potential it is easy to show that the equation of motion is given by

$$m \frac{d^2 \Delta x}{dt^2} + \left(\frac{\partial^2 W}{\partial x^2} \right) \Delta x = 0, \quad (6.7.2)$$

which is the simple harmonic oscillator equation. Consider what happens if the particle is given a small displacement at $t = 0$. If $\partial^2 W / \partial x^2 > 0$, the particle executes a simple harmonic oscillation and, if there is even a small amount of damping, the system eventually returns to the equilibrium point. The system is stable. On the other hand, if $\partial^2 W / \partial x^2 < 0$, the displacement grows exponentially and the particle never returns to the equilibrium point. The system is then unstable.

If the potential energy depends on n coordinates (x_1, x_2, \dots, x_n) , then static equilibria exist at points that simultaneously satisfy the n equations

$$\frac{\partial W(x_1, x_2, \dots, x_n)}{\partial x_\alpha} = 0, \quad \alpha = 1, 2, \dots, n. \quad (6.7.3)$$

Following the same type of analysis as above, it is easy to see that a given equilibrium point is stable if and only if for all α ,

$$\frac{\partial^2 W}{\partial x_\alpha^2} > 0, \quad \alpha = 1, 2, \dots, n. \quad (6.7.4)$$

For a system with a very large number of degrees of freedom, such as an MHD fluid, there are so many conditions required for the system to be stable that a completely stable equilibrium is more the exception than the rule.

6.7.2 Interchange instabilities

Although a wide variety of instabilities can occur in an MHD system, we start by focusing on a specific class of instabilities called interchange instabilities. Interchange instabilities involve a specific type of perturbation, namely the interchange of plasma between two adjacent magnetic flux tubes. Since the frozen field theorem applies, the magnetic flux

must be conserved in this interchange process. From equation (6.4.18) for the potential energy, W , one can see that if such a perturbation, δ , is applied to the system, the resulting change in the total energy consists of two terms

$$\delta W = \delta \int_V \frac{P}{\gamma - 1} d^3x + \delta \int_V \frac{B^2}{2\mu_0} d^3x, \quad (6.7.5)$$

where the first term represents the change in the internal energy, and the second term represents the change in the magnetic energy (see Section 6.4). To simplify the analysis further, we assume that the change in the internal energy is zero, so that the change in the potential energy is controlled entirely by the change in the magnetic energy. Although restrictive, this simple model can be used to show that a number of important MHD equilibria are unstable. The disadvantage is that we cannot show that a system is stable, since we have not considered all possible perturbations and the possible destabilizing effect of the internal energy. To determine whether a system is stable in all circumstances we need some more powerful tools, which will be discussed later in this chapter.

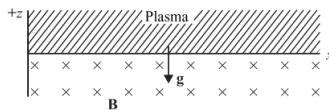


Fig. 6.21 A geometry in which an MHD fluid in a gravitational field \mathbf{g} is supported by a magnetic field \mathbf{B} .

As a simple example, consider an MHD fluid that is supported against gravity by a uniform horizontal magnetic field as shown in Figure 6.21. The static equilibrium condition for this geometry is simply that the gravitation force per unit area acting on the fluid must be balanced by the magnetic pressure on the underside of the fluid,

$$\frac{B^2}{2\mu_0} = \int_0^\infty \rho_m g dz. \quad (6.7.6)$$

The magnetic force arises because of a current \mathbf{J} that must flow along the boundary between the plasma and the magnetic field. By adding a gravitational potential energy term, $\rho_m g z$, to the potential energy equation (6.4.18), it is easy to see that the change in the potential energy of the system is given by

$$\delta W = \delta \int_V \frac{B^2}{2\mu_0} d^3x + \delta \int_V \rho_m g z d^3x. \quad (6.7.7)$$

Suppose now that a tube of fluid is interchanged with a magnetic flux tube of the same volume as shown in Figure 6.22. Since the magnetic field lines are straight and the volume elements before and after the interchange are the same, this perturbation does not change the magnetic energy. However, it does cause a change in the gravitational potential energy. The change in the potential energy of the system is given by

$$\delta W = \int_V \rho_m g \delta z d^3x, \quad (6.7.8)$$

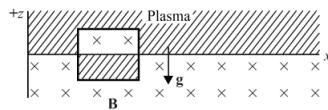


Fig. 6.22 An example of an interchange motion that replaces a magnetic flux tube with a tube of fluid from the other side of the boundary.

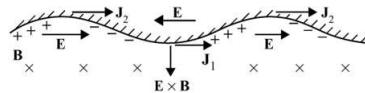


Fig. 6.23 When a perturbation develops, electrical currents cause charges to accumulate along the boundary. These charges lead to electric fields and $\mathbf{E} \times \mathbf{B}$ drifts that amplify the original perturbation.

which is negative, since the vertical displacement of the fluid element, δz , is negative. Thus, since the potential energy of the system has decreased, it follows that the system is unstable. This instability, which was first analyzed in detail by Kruskal and Schwarzschild [1954], is the MHD analog of the well-known Rayleigh–Taylor instability in fluid dynamics [see Batchelor, 1967].

The physical origin of the Kruskal–Schwarzschild instability can be seen by considering the $\mathbf{J} \times \mathbf{B}$ forces that occur at the boundary between the plasma and magnetic field. As shown in Figure 6.23, if a ripple-like perturbation develops in the boundary, the current at the low point, \mathbf{J}_1 , must be larger than the current at the high points, \mathbf{J}_2 , because more plasma is being supported against the gravitational force. The resulting difference in the boundary current causes regions of positive and negative charge to develop along the boundary, thereby producing oppositely directed \mathbf{E} fields at the high and low points of the ripple. The electric fields are directed in such a way that the $\mathbf{E} \times \mathbf{B}$ drifts reinforce the original disturbance, thereby causing the ripple to grow in amplitude.

Similar types of interchange motions also occur in plasmas that have large inertial forces caused by curvature in the plasma flow. For example, strong radial electric fields often develop in cylindrically symmetric plasma devices. As discussed earlier, these electric fields cause the plasma to rotate rapidly about the axis of the device. The centrifugal force produced by this rotation plays the same role as gravity, and leads to the development of interchange motions on the outer boundary of the plasma. The resulting perturbations look like the flutes on a Greek column, as shown in Figure 6.24. For this reason the instability is sometimes called a “flute” instability.

Interchange instabilities can also occur in the absence of gravity if the plasma is confined within a curved boundary by a magnetic field. Consider the geometry shown in Figure 6.25. To test the stability of this static equilibrium, we compute the change in the potential

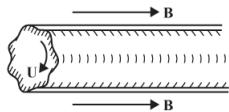


Fig. 6.24 In a rapidly rotating plasma column, the centrifugal force can drive an interchange instability, which leads to the development of flute-like ripples on the boundary of the column.

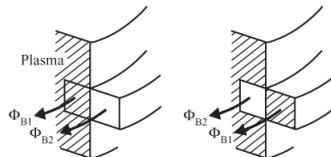


Fig. 6.25 An interchange motion consists of interchanging two flux tubes Φ_{B1} and Φ_{B2} at an interface between two plasmas.

energy, δW , when two adjacent flux tubes are interchanged while conserving the magnetic flux, Φ_{B1} and Φ_{B2} . The change in the potential energy is then given by

$$\delta W = \delta \int_V \frac{B^2}{2\mu_0} A \, d\ell = \delta \frac{\Phi_B^2}{2\mu_0} \int_V \frac{d\ell}{A}, \quad (6.7.9)$$

where A is the cross-sectional area of the flux tube and $d\ell$ is the distance along the magnetic field line. The change in magnetic energy of the system due to the interchange is then given by

$$\begin{aligned} \delta W &= \frac{1}{2\mu_0} \left[\Phi_{B2}^2 \left(\int_V \frac{d\ell_1}{A_1} - \int_V \frac{d\ell_2}{A_2} \right) + \Phi_{B1}^2 \left(\int_V \frac{d\ell_2}{A_2} - \int_V \frac{d\ell_1}{A_1} \right) \right] \\ &= \frac{1}{2\mu_0} (\Phi_{B2}^2 - \Phi_{B1}^2) \left(\int_V \frac{d\ell_1}{A_1} - \int_V \frac{d\ell_2}{A_2} \right) \\ &= \frac{1}{2\mu_0} (\Phi_{B2}^2 - \Phi_{B1}^2) \left(\int_V \left[1 - \frac{A_1}{A_2} \frac{d\ell_2}{d\ell_1} \right] \frac{d\ell_1}{A_1} \right). \end{aligned} \quad (6.7.10)$$

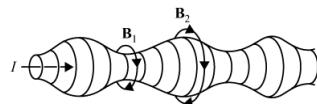


Fig. 6.26 The sausage-mode instability in a Z-pinch.

If the fluid is incompressible, which is assumed to be the case, we have $A_1 \, d\ell_1 = A_2 \, d\ell_2$, which gives

$$\delta W = \frac{1}{2\mu_0} (\Phi_{B2}^2 - \Phi_{B1}^2) \left\{ \int_V \left[1 - \left(\frac{d\ell_2}{d\ell_1} \right)^2 \right] \frac{d\ell_1}{A_1} \right\}. \quad (6.7.11)$$

Since the magnetic field must be stronger outside the fluid (to maintain the pressure balance), $\Phi_{B2} > \Phi_{B1}$. The above equation shows that if the field lines are concave towards the fluid ($d\ell_2 > d\ell_1$) the system is unstable. For example, the Z-pinch configuration is unstable to interchange motions. This instability manifests itself by the formation of “sausage-shaped” deformations in the outer surface of the plasma column as shown in Figure 6.26. This sausage-mode instability is essentially geometric in origin. Because the total current, I , that flows along the column must be constant, Ampère’s law, $B = \mu_0 I / (2\pi\rho)$, shows that the magnetic field increases in a region of smaller radius, which acts to increase the external magnetic pressure and thereby further shrinks the radius. Once the instability starts, the pinch is rapidly disrupted. Lightning discharges sometimes display this effect (called bead lightning). A Z-pinch can be stabilized by adding a magnetic field along the z axis. The axial magnetic field produces an internal magnetic field pressure in the region of smaller radius, thereby stabilizing the configuration. Lightning discharges also tend to be stabilized by atmospheric pressure. Therefore, only very intense discharges develop the bead-like structure.

Interchange instabilities also occur in the magnetospheres of Earth and Jupiter. In both cases, relatively low β plasmas (the plasmasphere at Earth and the Io plasma torus at Jupiter) occur deep within the magnetosphere, as shown in Figure 6.27.

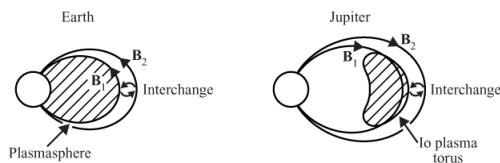


Fig. 6.27 The interchange instability is believed to play a major role in the radial transport of plasma in the magnetospheres of Earth and Jupiter.

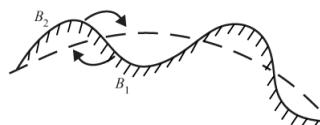


Fig. 6.28 The interchange instability tends to form flute-like ripples along the boundary of the plasma.

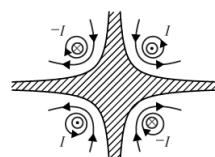


Fig. 6.29 The cusp mirror geometry.

Since the magnetic field lines in both cases are concave inward toward the region of higher plasma density, the plasmas tend to be unstable to interchange motions. The interchange

motions form flute-like ripples aligned parallel to the magnetic field as shown in Figure 6.28. Eventually the peaks in the ripple evolve into outward moving plumes that transport plasma away from the planet, and the valleys evolve into inward moving plumes that transport lower density plasma toward the planet. Interchange motions of this type can be stabilized by a variety of effects, one of which is the closure of magnetic field aligned currents through a resistive layer in the ionosphere. A full discussion of these stabilizing effects is beyond the scope of this book.

Since interchange instabilities of the type described above have their origin in the curvature of magnetic field lines, it is interesting to see what can be done to avoid these instabilities. This issue is particularly important for laboratory magnetic confinement devices, which attempt to confine a hot plasma with an external magnetic field. A simple magnetic mirror field, such as discussed in Section ??, tends to be unstable to interchange motions, since the plasma boundary is concave toward the central axis of the machine. One way of stabilizing such a device against interchange motions is to have the magnetic field lines concave outward away from the plasma. An example of such a configuration is the cusp mirror geometry. In the cusp mirror geometry, the plasma is confined in the low magnetic field region formed by four or more symmetrically arranged conductors with currents alternately in opposite directions, as shown in Figure 6.29. Since the boundaries of the plasma are concave outward away from the plasma, this geometry tends to stabilize the system against interchange motions.

6.7.3 The linear force operator

Although the above discussion illustrates some of the basic considerations involved in MHD stability, it is rather specialized. To take into account the effects of plasma pressure, current, and more general perturbations in inhomogeneous plasmas, we need to develop some more powerful tools for analyzing the change in the potential energy. Our immediate objective will be to re-express the forces on the right-hand side of the momentum equation (6.1.33) in the form of a linear operator, called the “linear force operator.” The linear force operator can then be used to evaluate the change in the potential energy for any arbitrary perturbation. As in previous analyses involving small perturbations, we will express all of the relevant MHD quantities as the sum of a time independent zero-order (0) quantity plus a small first-order perturbation. The time independent zero-order quantities, such as \mathbf{B}_0 , \mathbf{J}_0 , and P_0 , are assumed to obey the equations for static equilibrium, i.e., equations (6.6.6), (6.6.7), and (6.6.8). Note that, in contrast to previous perturbation analyses of this type, we cannot assume that the zero-order quantities are homogeneous. Unfortunately, this greatly increases the complexity of the analysis.

Since it is desirable to work with displacements rather than velocities, we introduce the perturbed fluid displacement vector $\xi(\mathbf{r}, t)$ by means of the equation

$$\mathbf{U} = \frac{\partial \xi(\mathbf{r}, t)}{\partial t}. \quad (6.7.12)$$

Using this relation, it is then easy to show that the linearized version of the momentum equation (6.1.33) is given by

$$\rho_{m0} \frac{\partial^2 \xi}{\partial t^2} = \frac{1}{\mu_0} [(\nabla \times \mathbf{B}) \times \mathbf{B}_0 + (\nabla \times \mathbf{B}_0) \times \mathbf{B}] - \nabla P, \quad (6.7.13)$$

where the current density has been eliminated using Ampère's law (6.1.28) and, as usual, we have omitted the subscript 1 on the first-order perturbed quantities. The induction equation (6.3.7) can be integrated over time to give

$$\mathbf{B} = \nabla \times (\xi \times \mathbf{B}_0). \quad (6.7.14)$$

To obtain an equation for the pressure we combine the linearized form of the mass continuity equation (6.1.32),

$$\frac{\partial \rho_m}{\partial t} + \rho_{m0} \nabla \cdot \mathbf{U} + \mathbf{U} \cdot \nabla \rho_{m0} = 0, \quad (6.7.15)$$

with the linearized form of the adiabatic equation of state (6.1.35),

$$\frac{\partial P}{\partial t} + \mathbf{U} \cdot \nabla P_0 - \frac{\gamma P_0}{\rho_{m0}} \frac{\partial \rho_m}{\partial t} - \frac{\gamma P_0}{\rho_{m0}} \mathbf{U} \cdot \nabla \rho_{m0} = 0, \quad (6.7.16)$$

to obtain

$$\frac{\partial P}{\partial t} + \mathbf{U} \cdot \nabla P_0 + \gamma P_0 (\nabla \cdot \mathbf{U}) = 0. \quad (6.7.17)$$

Replacing \mathbf{U} by $\partial \xi / \partial t$, and integrating over time gives

$$P = -\xi \cdot \nabla P_0 - \gamma P_0 \nabla \cdot \xi. \quad (6.7.18)$$

Substituting expressions (6.7.14) and (6.7.18) for \mathbf{B} and P , respectively, into equation (6.7.13), we obtain the linear differential equation

$$\rho_{m0} \frac{\partial^2 \xi}{\partial t^2} = \mathbf{F}(\xi), \quad (6.7.19)$$

where $\mathbf{F}(\xi)$, the linear force operator, is given by

$$\begin{aligned} \mathbf{F}(\xi) &= \frac{1}{\mu_0} [(\nabla \times \{\nabla \times (\xi \times \mathbf{B}_0)\}) \times \mathbf{B}_0 + (\nabla \times \mathbf{B}_0) \times \{\nabla \times (\xi \times \mathbf{B}_0)\}] \\ &\quad + \nabla [\xi \cdot \nabla P_0 + \gamma P_0 (\nabla \cdot \xi)]. \end{aligned} \quad (6.7.20)$$

In principle, given an arbitrary static equilibrium, the equation of motion (6.7.19) can then be solved to determine all possible plasma displacements $\xi(\mathbf{r}, t)$. In practice, because of the complexity of the linear force operator, such an approach is difficult (even on a computer) for all but the simplest cases. However, the linear force operator has some very nice mathematical properties that greatly simplify the analysis, and lead to two useful and sometimes complementary approaches, called (1) the normal mode method and (2) the energy method. We next describe these two methods.

6.7.4 The normal mode method

In the normal mode method, we assume that the displacement vector can be written as a simple harmonic function of the form

$$\xi_n(\mathbf{r}, t) = \xi_n(\mathbf{r}) \exp(-i\omega_n t), \quad (6.7.21)$$

where the subscript n designates a particular normal mode of frequency ω_n . Since the equation of motion (6.7.19) is linear in ξ , we can construct a general solution $\xi(\mathbf{r}, t)$ by a linear superposition, i.e.,

$$\xi(\mathbf{r}, t) = \sum_n \xi_n(\mathbf{r}) \exp(-i\omega_n t). \quad (6.7.22)$$

To find the constraints on the normal mode frequencies, we start by substituting expression (6.7.21) into equation (6.7.19), which gives

$$-\rho_{m0} \omega_n^2 \xi_n = \mathbf{F}(\xi_n). \quad (6.7.23)$$

Next, we state the following theorem, the proof of which is assigned as an exercise.

Theorem The linear force operator \mathbf{F} is self-adjoint, i.e., if there are two eigenfunctions $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ (satisfying standard boundary conditions), then

$$\int_V \boldsymbol{\eta}_2 \cdot \mathbf{F}(\boldsymbol{\eta}_1) d^3x = \int_V \boldsymbol{\eta}_1 \cdot \mathbf{F}(\boldsymbol{\eta}_2) d^3x. \quad (6.7.24)$$

This theorem has the following corollaries.

Corollary 1 ω_n^2 is always real.

To prove Corollary 1, consider the complex conjugate of equation (6.7.23),

$$-\rho_{m0} \omega_n^{2*} \xi_n^* = \mathbf{F}(\xi_n^*). \quad (6.7.25)$$

Taking the scalar product of both sides of the above equation with ξ_n and integrating over the entire volume of the system, gives

$$-\rho_{m0} \omega_n^{2*} \int_V \xi_n \cdot \xi_n^* d^3x = \int_V \xi_n \cdot \mathbf{F}(\xi_n^*) d^3x. \quad (6.7.26)$$

Identifying $\boldsymbol{\eta}_1 = \xi_n$ and $\boldsymbol{\eta}_2 = \xi_n^*$, we can then invoke equation (6.7.24) and rewrite the above equation as

$$\begin{aligned} -\rho_{m0}\omega_n^{2*} \int_V \boldsymbol{\xi}_n \cdot \boldsymbol{\xi}_n^* d^3x &= \int_V \boldsymbol{\xi}_n^* \cdot F(\boldsymbol{\xi}_n) d^3x \\ &= -\rho_{m0}\omega_n^2 \int_V \boldsymbol{\xi}_n^* \cdot \boldsymbol{\xi}_n d^3x. \end{aligned} \quad (6.7.27)$$

Since $\int \boldsymbol{\xi}_n \cdot \boldsymbol{\xi}_n^* d^3x = \int |\boldsymbol{\xi}_n|^2 d^3x$ is non-zero for any non-trivial eigenfunction, it follows from the above equation that

$$\omega_n^2 = \omega_n^{2*}. \quad (6.7.28)$$

Therefore, ω_n^2 is always real. This result implies that the eigenfrequency ω_n is either purely real or purely imaginary. If ω_n is imaginary, then the system is unstable, since one of the two roots leads to exponential growth in equation (6.7.21). Note that, if the self-adjointness of \mathbf{F} is destroyed, as in an equilibrium with flows, i.e., $\mathbf{U}_0 \neq 0$, then equation (6.7.28) does not hold.

Corollary 2 The eigenmodes of \mathbf{F} are orthonormal, i.e., they obey the condition

$$\int_V \rho_{m0} \boldsymbol{\xi}_m^* \cdot \boldsymbol{\xi}_n d^3x = \delta_{mn}, \quad (6.7.29)$$

where δ_{mn} is the Kronecker delta function ($\delta_{mn} = 1$ if $m = n$, or 0 if $m \neq n$).

To prove Corollary 2, consider the expression

$$\begin{aligned} (\omega_m^{2*} - \omega_n^2) \int_V \rho_{m0} \boldsymbol{\xi}_m^* \cdot \boldsymbol{\xi}_n d^3x \\ = - \int_V [\boldsymbol{\xi}_n \cdot \mathbf{F}(\boldsymbol{\xi}_m^*) - \boldsymbol{\xi}_m^* \cdot \mathbf{F}(\boldsymbol{\xi}_n)] d^3x \\ = 0. \end{aligned} \quad (6.7.30)$$

The last step follows from equation (6.7.24). If $m \neq n$, equation (6.7.30) implies that

$$\int_V \rho_{m0} \boldsymbol{\xi}_m^* \cdot \boldsymbol{\xi}_n d^3x = 0. \quad (6.7.31)$$

If $m = n$, equation (6.7.30) is satisfied automatically because of relation (6.7.28). In that case, one can normalize the linear eigenfunction $\boldsymbol{\xi}$ so that

$$\int_V \rho_{m0} \boldsymbol{\xi}_m^* \cdot \boldsymbol{\xi}_m d^3x = 1. \quad (6.7.32)$$

Combining equations (6.7.31) and (6.7.32), we obtain equation (6.7.29).

As mentioned above, an important consequence of Corollary 1 is that, for an ideal MHD static equilibrium, the eigenfrequency is either purely oscillatory (i.e., ω_n is purely real) or purely growing (i.e., ω_n is purely imaginary). This makes the marginal frequency condition, $\omega^2 = 0$, an important case to study, since it separates stable solutions from unstable

solutions. Thus, Corollary 1 is of great help in determining the stability of the system, since one need not search the entire complex ω^2 plane but only the real ω^2 axis for solutions, with the origin $\omega^2 = 0$ separating unstable solutions ($\omega^2 < 0$) from stable ones ($\omega^2 > 0$). Schematically, a possible frequency spectrum of the ideal MHD force operator is shown in Figure 6.30. The origin 0 is the marginal point. To the left of the origin are discrete and unstable frequencies. To the right of the origin are stable frequencies. It is possible to obtain not only discrete stable frequencies, but also a continuous band of stable frequencies, indicated by the dark shaded band.

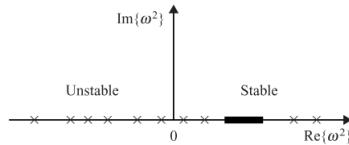


Fig. 6.30 The frequency spectrum of the ideal MHD force operator.

The normal mode method is a brute force method for determining the eigenmodes and eigenfrequencies of the system using equation (6.7.23). Once these are known, the general solution (6.7.22) can be constructed by superposition. Next, we discuss the energy principle, which is a more subtle and even more powerful method of testing the stability of an MHD system.

6.7.5 The energy principle

To analyze the stability of a static MHD equilibrium using the energy principle, we essentially follow the same procedure discussed in Section 6.7.1 for a particle moving in a one-dimensional potential. However, the analysis is more complicated, since an MHD fluid has an infinite number of degrees of freedom. The first step is to compute the change in the potential energy, δW , for a small arbitrary perturbation, ξ , relative to some equilibrium state. To obtain the change in the potential energy, δW , we start by showing that

$$\frac{d}{dt} \int_V \left[\frac{1}{2} \rho_{m0} \left| \frac{\partial \xi}{\partial t} \right|^2 - \frac{1}{2} \xi^* \cdot \mathbf{F}(\xi) \right] d^3x = 0, \quad (6.7.33)$$

where $\mathbf{F}(\xi)$ is the linear force operator defined by equation (6.7.20). To prove the above result, consider the left-hand side, which can be written

$$\begin{aligned} & \frac{1}{2} \int_V \left[\rho_{m0} \frac{\partial \xi^*}{\partial t} \cdot \frac{\partial^2 \xi}{\partial t^2} + \rho_{m0} \frac{\partial \xi}{\partial t} \cdot \frac{\partial^2 \xi^*}{\partial t^2} - \frac{\partial \xi^*}{\partial t} \cdot \mathbf{F}(\xi) - \xi^* \cdot \mathbf{F}\left(\frac{\partial \xi}{\partial t}\right) \right] d^3x \\ &= \frac{1}{2} \int_V \left[\frac{\partial \xi^*}{\partial t} \cdot \left\{ \rho_{m0} \frac{\partial^2 \xi}{\partial t^2} - \mathbf{F}(\xi) \right\} + \frac{\partial \xi}{\partial t} \cdot \left\{ \rho_{m0} \frac{\partial^2 \xi^*}{\partial t^2} - \mathbf{F}(\xi^*) \right\} \right] d^3x, \end{aligned} \quad (6.7.34)$$

where in the last step we have used the equation (6.7.24) to write

$$\int_V \boldsymbol{\xi}^* \cdot \mathbf{F} \left(\frac{\partial \boldsymbol{\xi}}{\partial t} \right) d^3x = \int_V \frac{\partial \boldsymbol{\xi}}{\partial t} \cdot \mathbf{F}(\boldsymbol{\xi}^*) d^3x. \quad (6.7.35)$$

It is obvious now that, upon using the linear force equation (6.7.19), expression (6.7.34) vanishes identically.

By simple inspection, one can see that equation (6.7.33) is an energy conservation equation for the perturbed system. It implies that

$$\delta K + \delta W = C = \text{constant}, \quad (6.7.36)$$

where

$$\delta K = \frac{1}{2} \int_V \rho_{m0} \left| \frac{\partial \boldsymbol{\xi}}{\partial t} \right|^2 d^3x \quad (6.7.37)$$

is the perturbed kinetic energy and

$$\delta W = -\frac{1}{2} \int_V \boldsymbol{\xi}^* \cdot \mathbf{F}(\boldsymbol{\xi}) d^3x \quad (6.7.38)$$

is the perturbed potential energy. Note that this equation differs from equation (6.4.18) for the total energy, W , because it represents the energy change, δW , relative to the equilibrium state, due to the linearized small amplitude perturbation.

We now prove that $\delta W \geq 0$ is a necessary and sufficient condition for stability. If a fluid displacement is unstable, the kinetic energy δK will increase in time without bound. (It will increase exponentially for a linear instability.) But while δK increases, δW will have to decrease and become negative in order to keep the energy C given by equation (6.7.36) constant. In contrast, if $\delta W \geq 0$, δK must be bounded from above according to the relation $\delta K \leq C - \delta W$. In this case, there can be no unbounded growth of kinetic energy and the plasma is stable. Hence, $\delta W \geq 0$ is a sufficient condition for stability.

To prove that $\delta W \geq 0$ is also a necessary condition for stability, let us consider an initial displacement such that $\boldsymbol{\xi}(\mathbf{r}, 0) \neq 0$ but $\partial \boldsymbol{\xi}(\mathbf{r}, 0)/\partial t = 0$. Furthermore, we require that this initial displacement makes δW negative. It follows that

$$\begin{aligned} C(t=0) &= \delta W(t=0) + \delta K(t=0) \\ &= \delta W(t=0) < 0. \end{aligned} \quad (6.7.39)$$

Let us define the quantity

$$I(t) = \frac{1}{2} \int_V \rho_{m0} |\boldsymbol{\xi}|^2 d^3x. \quad (6.7.40)$$

The first derivative of I is given by

$$\frac{dI}{dt} = \frac{1}{2} \int_V \rho_{m0} \left(\boldsymbol{\xi}^* \cdot \frac{\partial \boldsymbol{\xi}}{\partial t} + \boldsymbol{\xi} \cdot \frac{\partial \boldsymbol{\xi}^*}{\partial t} \right) d^3x, \quad (6.7.41)$$

and the second derivative is given by

$$\begin{aligned}\frac{d^2I}{dt^2} &= \frac{1}{2} \int_V \rho_{m0} \left(2 \left| \frac{\partial \xi}{\partial t} \right|^2 + \xi^* \cdot \frac{\partial^2 \xi}{\partial t^2} + \xi \cdot \frac{\partial^2 \xi^*}{\partial t^2} \right) d^3x \\ &= \frac{1}{2} \int_V \left[2 \rho_{m0} \left| \frac{\partial \xi}{\partial t} \right|^2 + \xi^* \cdot \mathbf{F}(\xi) + \xi \cdot \mathbf{F}(\xi^*) \right] d^3x.\end{aligned}\quad (6.7.42)$$

Finally, using equations (6.7.24), (6.7.37) and (6.7.38), we obtain

$$\begin{aligned}\frac{d^2I}{dt^2} &= 2(\delta K - \delta W) \\ &= 2(2\delta K - C),\end{aligned}\quad (6.7.43)$$

where the last step follows from equation (6.7.36). Since δK is positive definite at all times and C is negative by virtue of equation (6.7.39), we can write

$$\frac{d^2I}{dt^2} > -2C > 0.\quad (6.7.44)$$

The above equation implies that I , and hence ξ , increases without bound if there exists any initial perturbation that makes δW negative. To insure stability, $\delta W \geq 0$ must hold for all permissible plasma displacements. This proves the necessity of the stability condition $\delta W \geq 0$.

In addition to the strength of the energy principle, $\delta W \geq 0$, as a criterion for ideal MHD stability, there is yet another property of δW that makes it a very useful computational tool. It can be shown that δW is variational [Freidberg, 1987]. By this we mean the following: consider a trial function ξ that obeys the boundary conditions that a true plasma displacement of the system must obey, although ξ itself may not be a true displacement. If this trial function makes δW negative, then the variational property of δW guarantees that there exists a true displacement ξ that makes δW even more negative.

The variational property of δW enables us to settle rigorously the stability or instability of any given magnetostatic equilibrium without knowledge of the true eigenfunctions of the system. All that we need to do is to have some intuition regarding a trial function obeying the appropriate boundary conditions that can make δW negative, and that is quite enough to demonstrate the equilibrium is unstable, because a true eigenfunction will make δW even more negative. We state this property without proof. See Freidberg [1987] for a further discussion of this topic.

6.7.6 A more useful form for δW

Equation (6.7.38) for the perturbed potential is often not in a convenient form for calculations. A more useful form can be obtained by substituting equation (6.7.20) for the linear force operator, $\mathbf{F}(\xi)$, into equation (6.7.38) and simplifying the resulting equation. After carrying out this substitution, equation (6.7.38) becomes

$$\delta W = -\frac{1}{2} \int_V \boldsymbol{\xi}^* \cdot \left\{ \frac{1}{\mu_0} [(\nabla \times \{\nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0)\}) \times \mathbf{B}_0 + (\nabla \times \mathbf{B}_0) \times \{\nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0)\}] + \nabla[\boldsymbol{\xi} \cdot \nabla P_0 + \gamma P_0 (\nabla \cdot \boldsymbol{\xi})] \right\} d^3x. \quad (6.7.45)$$

To proceed further, we assume that the plasma is surrounded by a perfectly conducting wall. In this case, the plasma displacement $\boldsymbol{\xi}$ as well as the perturbed magnetic field \mathbf{B} , given by equation (6.7.14), obey the boundary conditions

$$\hat{\mathbf{n}} \cdot \boldsymbol{\xi} = 0 \quad (6.7.46)$$

and

$$\hat{\mathbf{n}} \cdot \mathbf{B}_1 = \hat{\mathbf{n}} \cdot [\nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0)] = 0. \quad (6.7.47)$$

We next consider the terms in the integrand on the right of equation (6.7.45). Using the vector identity $\nabla \cdot (\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot (\nabla \times \mathbf{F}) - \mathbf{F} \cdot (\nabla \times \mathbf{G})$, the first term in the integrand can be written

$$\begin{aligned} & \frac{1}{\mu_0} \boldsymbol{\xi}^* \cdot \{(\nabla \times \mathbf{B}_1) \times \mathbf{B}_0\} \\ &= \frac{1}{\mu_0} (\mathbf{B}_0 \times \boldsymbol{\xi}^*) \cdot (\nabla \times \mathbf{B}_1) \\ &= -\frac{1}{\mu_0} \nabla \cdot \{(\mathbf{B}_0 \times \boldsymbol{\xi}^*) \times \mathbf{B}_1\} + \frac{1}{\mu_0} \mathbf{B}_1 \cdot \nabla \times (\mathbf{B}_0 \times \boldsymbol{\xi}^*) \\ &= \frac{1}{\mu_0} \nabla \cdot \{(\boldsymbol{\xi}^* \times \mathbf{B}_0) \times \mathbf{B}_1\} - \frac{1}{\mu_0} |\nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0)|^2. \end{aligned} \quad (6.7.48)$$

The second term in the integrand can be simply rewritten as

$$\frac{1}{\mu_0} \boldsymbol{\xi}^* \cdot (\nabla \times \mathbf{B}_0) \times \{\nabla \times (\boldsymbol{\xi} \times \mathbf{B})\} = \boldsymbol{\xi}^* \cdot \mathbf{J}_0 \times \{\nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0)\}. \quad (6.7.49)$$

The third term needs no reprocessing at this point, so we move on to the fourth term, which can be written

$$\begin{aligned} \boldsymbol{\xi}^* \nabla \{\gamma P_0 (\nabla \cdot \boldsymbol{\xi})\} &= \nabla \cdot \{\gamma P_0 \boldsymbol{\xi}^* (\nabla \cdot \boldsymbol{\xi})\} - \gamma P_0 (\nabla \cdot \boldsymbol{\xi}^*) (\nabla \cdot \boldsymbol{\xi}) \\ &= \nabla \cdot \{\gamma P_0 \boldsymbol{\xi}^* (\nabla \cdot \boldsymbol{\xi})\} - \gamma P_0 |\nabla \cdot \boldsymbol{\xi}|^2. \end{aligned} \quad (6.7.50)$$

Substituting equations (6.7.48), (6.7.49), and (6.7.50) into expression (6.7.45) for δW , and using Gauss' theorem to convert the divergence terms in equations (6.7.48) and (6.7.50) to surface terms, we obtain

$$\begin{aligned}\delta W = & \frac{1}{2} \int_V \left[\frac{1}{\mu_0} \|\nabla \times (\xi \times \mathbf{B}_0)\|^2 + \gamma P_0 |\nabla \cdot \xi|^2 \right. \\ & \left. - \xi^* \cdot \mathbf{J}_0 \times \{\nabla \times (\xi \times \mathbf{B}_0)\} - \xi^* \cdot \nabla(\xi \cdot \nabla P_0) \right] d^3x \\ & - \frac{1}{2} \int_S \hat{\mathbf{n}} \cdot \left[\frac{1}{\mu_0} ((\xi^* \times \mathbf{B}_0) \times \mathbf{B}_1) + \gamma P_0 \xi^* (\nabla \cdot \xi) \right] dA,\end{aligned}\quad (6.7.51)$$

where $\hat{\mathbf{n}}$ is the unit normal on the surface S bounding the plasma. When S is perfectly conducting, the surface integral vanishes because of boundary conditions (6.7.46) and (6.7.47), and we are left with

$$\begin{aligned}\delta W = & \frac{1}{2} \int_V \left[\frac{1}{\mu_0} \|\nabla \times (\xi \times \mathbf{B}_0)\|^2 + \gamma P_0 |\nabla \cdot \xi|^2 \right. \\ & \left. - \xi^* \cdot \mathbf{J}_0 \times \{\nabla \times (\xi \times \mathbf{B}_0)\} - \xi^* \cdot \nabla(\xi \cdot \nabla P_0) \right] d^3x.\end{aligned}\quad (6.7.52)$$

The first two terms on the right-hand side of the above equation are positive and stabilizing, while the third and fourth terms can be destabilizing. The first term represents the energy required to bend field lines. The second term represents the energy required to compress a plasma with non-zero equilibrium pressure. The third term, which depends explicitly on the equilibrium current density, \mathbf{J}_0 , can potentially drive instabilities of the “kink” type. The fourth term, which depends explicitly on the equilibrium pressure gradient, can potentially drive instabilities of the “ballooning” or “interchange” type. For an ideal MHD plasma, an instability always arranges its eigenfunction in such a way as to minimize the stabilizing contributions to δW (the first and second terms). For example, in an infinite cylinder or a torus, the marginally stable ($\omega_n = 0$) eigenfunctions are incompressible and obey the condition $\nabla \cdot \xi = 0$, which reduces the second stabilizing term in δW to zero.

The current density and the pressure gradient in any given magnetostatic equilibrium are potential sources of free energy for ideal MHD instabilities. In most cases of physical interest, δW needs to be computed numerically to determine the stability for various perturbations, such as kink, ballooning, and interchange motions. In the next subsection, we consider an analytically tractable example.

6.7.7 The Rayleigh-Taylor instability

We next discuss the well-known instability known as the Rayleigh-Taylor instability, which is similar to the Kruskal-Schwartzschild instability discussed earlier except, instead of an abrupt boundary, the vertical distribution of plasma density is continuous. To illustrate the techniques involved, we will carry out the analysis using both the normal mode method and the energy method. The basic geometry is shown Figure 6.31. In this model we assume that the plasma is bounded by perfectly conducting walls at $z = 0$ and $z = a$, that the gravitational acceleration, \mathbf{g} , is directed downward in the $-z$ direction, and that the magnetic field, \mathbf{B}_0 , is in the $+y$ direction (i.e., into the paper). To provide for a simple vertical

variation of the fluid density, we assume that the equilibrium mass density profile is given by an exponential

$$\rho_{m0}(z) = \rho_0 \exp(-z/H_S) \quad (6.7.53)$$

where ρ_0 is the mass density at $z = 0$, and H_S is a constant that is called the scale height in the ionospheric literature. If $H_S > 0$, the equilibrium density decreases with increasing height, and if $H_S < 0$, the equilibrium density increases with increasing height. Because of the downward gravitational force, if $H_S < 0$ we expect the equilibrium to be unstable, since this represents a situation in which a heavy fluid lies on top of a light fluid.

In the presence of gravity, the momentum equation (6.2.10) must be modified to include the gravitational force, $\rho_m \mathbf{g}$, so that

$$\rho_m \frac{d\mathbf{U}}{dt} = \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B} - \nabla \left(\frac{B^2}{2\mu_0} + P \right) + \rho_m \mathbf{g}. \quad (6.7.54)$$

In magnetostatic equilibrium, i.e., $U_0 = 0$, the above equation with $\mathbf{B} = B_0 \hat{\mathbf{y}}$ reduces to

$$\frac{\partial}{\partial z} \left(\frac{B_0^2}{2\mu_0} + P_0 \right) + \rho_{m0} g = 0. \quad (6.7.55)$$

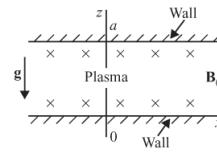


Fig. 6.31

An ideal MHD fluid between two perfectly conducting walls (at $z = 0$ and $z = a$) with a downward applied gravitational force, $\rho_m g$. The magnetic field, indicated by the \times symbols, is directed into the paper.

Linearizing this equation with respect to small perturbations around the zero-order equilibrium gives

$$\rho_{m0} \frac{\partial \mathbf{U}}{\partial t} = \frac{1}{\mu_0} (\mathbf{B}_0 \cdot \nabla) \mathbf{B} + \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B}_0 - \nabla \left(\frac{\mathbf{B}_0 \cdot \mathbf{B}}{\mu_0} + P \right) + \rho_m \mathbf{g}, \quad (6.7.56)$$

where, as usual, we have suppressed the subscript 1 on the first-order terms. Next consider a plane wave perturbation of the form $\exp(i\omega t + ikx)$. Note that in considering such a form, we have assumed that the perturbations have no spatial variation along y , and hence, the equilibrium magnetic field suffers no spatial modulation along y . Also, because the equilibrium depends explicitly on z , we do not assume plane wave solutions along z , but solve for the z dependence of the solutions from the linearized equations and boundary conditions. Under these conditions, the first two terms on the right of equation (6.7.56) vanish identically. It is convenient to eliminate the third term by taking the curl of both sides of equation (6.7.56), which gives

$$\nabla \times \left(\rho_{m0} \frac{\partial \mathbf{U}}{\partial t} \right) = \nabla \times (\rho_m \mathbf{g}). \quad (6.7.57)$$

The y component of the above equation is given by

$$\omega \left[ik\rho_{m0}\tilde{U}_z - \frac{\partial}{\partial z}(\rho_{m0}\tilde{U}_x) \right] = -kg\tilde{\rho}_m. \quad (6.7.58)$$

Next we assume, for simplicity, that the perturbed velocity is incompressible, i.e.,

$$\nabla \cdot \mathbf{U} = 0. \quad (6.7.59)$$

For the assumed plane wave perturbation, the incompressibility condition then becomes

$$ik\tilde{U}_x + \frac{\partial \tilde{U}_z}{\partial z} = 0. \quad (6.7.60)$$

Since $\nabla \cdot \mathbf{U} = 0$, the linearized mass continuity equation (6.7.15) simplifies to

$$\frac{\partial \rho_m}{\partial t} + \mathbf{U} \cdot \nabla \rho_{m0} = 0, \quad (6.7.61)$$

which for the assumed plane wave perturbation yields

$$i\omega\tilde{\rho}_m + \tilde{U}_z \frac{\partial \rho_{m0}}{\partial z} = 0. \quad (6.7.62)$$

It is straightforward to eliminate \hat{U}_x and $\hat{\rho}_{m0}$ from equation (6.7.58) by using equations (6.7.60) and (6.7.62), which gives

$$\frac{1}{\rho_{m0}} \frac{\partial}{\partial z} \left(\rho_{m0} \frac{\partial \tilde{U}_z}{\partial z} \right) - k^2 \left(1 - \frac{g}{H_S \omega^2} \right) \tilde{U}_z = 0, \quad (6.7.63)$$

where $1/H_S = -(1/\rho_{m0})\partial \rho_{m0}/\partial z$.

The above equation is a second-order differential equation for \tilde{U}_z , to be solved using the boundary conditions $\tilde{U}_z = 0$ at $z = 0$ and $z = a$. To solve equation (6.7.63) in closed form, we make the substitution

$$\tilde{U}_z(z) = f(z) \exp\left(\frac{z}{2H_S}\right), \quad (6.7.64)$$

which gives

$$\frac{d^2 f}{dz^2} + \alpha^2 f = 0 \quad (6.7.65)$$

where

$$\alpha^2 = k^2 \left(\frac{g}{H_S \omega^2} - 1 \right) - \frac{1}{4H_S^2}. \quad (6.7.66)$$

Since $f(0) = f(a) = 0$, we obtain the eigenfunctions

$$f_n(z) = f_0 \sin \frac{n\pi z}{a}, \quad n = 1, 2, \dots, \quad (6.7.67)$$

where f_0 is a constant, and the eigenvalue relation

$$\frac{n^2\pi^2}{a^2} = k^2 \left(\frac{g}{H_S \omega_n^2} - 1 \right) - \frac{1}{4H_S^2}. \quad (6.7.68)$$

The above equation can be solved for the frequencies of the normal modes which are given by

$$\omega_n^2 = \left(\frac{g}{H_S} \right) \frac{4k^2 a^2 H_S^2}{a^2 + 4H_S^2(K^2 a^2 + n^2 \pi^2)}. \quad (6.7.69)$$

Note that:

1. If $H_S > 0$, the frequency, ω_n , of the n th mode is purely real and the system is stable. On the other hand, if $H_S < 0$, the frequency is purely imaginary, and the system is unstable.
2. The largest growth rate, $\sqrt{|g/H_S|}$, occurs in the limit $k \rightarrow \infty$ and the smallest growth rate, occurs in the limit $k \rightarrow 0$.
3. For fixed k , the growth rate decreases with increasing mode number n .

The Rayleigh–Taylor instability can also be analyzed by means of the energy method. Note that in this case, we need to add the gravity term $\rho_m \mathbf{g}$ to the linear force operator $\mathbf{F}(\xi)$, equation (6.7.20). Assuming incompressibility, i.e., $\nabla \cdot \xi = 0$, we can write

$$\begin{aligned} \rho_m \mathbf{g} &= -\nabla \cdot (\rho_{m0}) \mathbf{g} \\ &= \xi_z (\partial \rho_{m0} / \partial z) g \hat{\mathbf{z}}. \end{aligned} \quad (6.7.70)$$

The contribution of the term $\rho_m \mathbf{g}$ to δW is given by $(-1/2) \xi^* \cdot \rho_m \mathbf{g} = -(1/2) |\xi_z|^2 g \partial \rho_{m0} / \partial z$. While this contribution needs to be added to the integrand in equation (6.7.52), note that the second and third terms in the integrand vanish because $\nabla \cdot \xi = 0$ and $\mathbf{J}_0 = 0$, respectively. Hence, we obtain

$$\delta W = \frac{1}{2} \int_V \left[\frac{1}{\mu_0} |\nabla \times (\xi \times \mathbf{B}_0)|^2 - (\xi^* \cdot \nabla P_0) - |\xi_z|^2 g \partial \rho_{m0} / \partial z \right] d^3 x. \quad (6.7.71)$$

Note that the second term in the integrand on the right-hand side of the above equation can be written as $\nabla \cdot [\xi^* (\xi \cdot \nabla P_0)]$, which can be transformed by Gauss' law to a surface term that vanishes upon using the boundary condition (6.7.46). This leaves the third term in the integrand as the only possible destabilizing term. Hence a sufficient condition for stability is

$$\frac{\partial \rho_{m0}}{\partial z} < 0 \quad \text{or} \quad H_S > 0, \quad (6.7.72)$$

which is consistent with the results determined from the normal mode method. A good example of the Rayleigh–Taylor instability occurs in the Earth’s ionosphere near the magnetic equator, where the magnetic field is nearly horizontal. Under certain conditions, “bubbles” of low density plasma from near the base of the ionosphere rise upwards into the ionosphere, causing plume-like disturbances, such as shown in Figure 6.32, that can be detected by ground-based radars [Basu and Kelley, 1979].

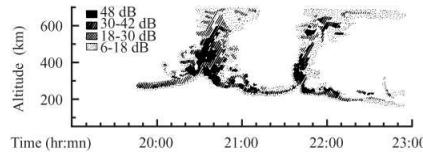


Fig. 6.32 A radar profile of the Earth’s ionosphere that shows buoyant rising plumes of plasma caused by the Rayleigh–Taylor instability [from Basu and Kelley, 1979].

6.8 Magnetic reconnection

In Section 3.6.3.1, we discussed particle motion in a basic x -type magnetic field configuration and the possibility of plasma transport across a boundary between regions of different magnetic field topologies. This type of plasma transport in two dimensions is called magnetic reconnection or field line merging [Vasyliunas, 1975]. In what follows, we give a brief discussion of some of the important features and models of this important phenomenon. A classic example of a magnetic field configuration that can support reconnection is the so-called Harris current sheet, shown in Figure 6.33, for which the magnetic field is given by

$$\mathbf{B} = B_0 \tanh\left(\frac{y}{d}\right) \hat{x} + B_T \hat{z}, \quad B_0 > 0, \quad B_T > 0. \quad (6.8.1)$$

Note that at the $y = 0$ surface, $B_x = 0$. For $y > 0$, the magnetic field in the x, y plane points in the $+x$ direction, while for $y < 0$, it points in the $-x$ direction. The neutral surface $y = 0$, represented by the dashed line in Figure 6.33, has a special geometric significance because it separates two regions of oppositely directed magnetic field.

In addition to the proper magnetic geometry, an example of which is given in the figure, magnetic reconnection requires some form of dissipation that violates the frozen field theorem, discussed in Section 6.3.1. If plasma is transported across a boundary between regions with different magnetic field topologies, as required in magnetic reconnection, it must be that the plasma is not frozen to the magnetic field, but slips with respect to the field. The frozen field theorem, which is deduced from the condition $\mathbf{E} + \mathbf{U} \times \mathbf{B} = 0$, requires that

$E_{\parallel} = 0$. It follows that a necessary condition for magnetic reconnection to occur is that the component of the electric field parallel to the magnetic field, denoted by E_{\parallel} , be non-zero. In a resistive plasma with the Ohm's law (6.1.34), this implies that

$$E_{\parallel} = \eta J_{\parallel} \neq 0. \quad (6.8.2)$$

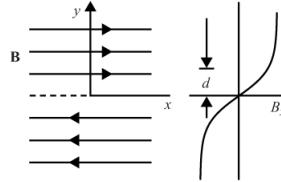


Fig. 6.33 The magnetic field geometry across a planar current sheet.

When $E_{\parallel} \neq 0$, the magnetic flux threading any closed curve moving with the plasma is not constant, violating Statement 1 of Section 6.3.1. Instead of flux conservation, we can now have flux annihilation.

Although necessary, the existence of a non-zero, parallel electric field, E_{\parallel} , is not sufficient to ensure that magnetic reconnection will occur. The change of magnetic flux in a closed loop due to resistive diffusion can produce a non-zero E_{\parallel} , but diffusion does not necessarily imply reconnection. One way to distinguish diffusion from reconnection is by considering characteristic time scales. In the absence of resistivity, the characteristic time scale of field line motion is the so-called Alfvén time scale

$$\tau_A = \frac{d}{V_A}, \quad (6.8.3)$$

where d is on the order of the system size, and $V_A = B_0 / \sqrt{\mu_0 \rho_m}$ is the characteristic Alfvén speed corresponding to the component B_x . The presence of resistivity introduces another time scale that can be obtained by balancing the term on the left of equation (6.3.4) with the second term on the right of equation (6.3.4). We then obtain the resistive diffusion time scale,

$$\tau_R = \mu_0 \sigma d^2 = \mu_0 \eta^{-1} d^2. \quad (6.8.4)$$

The Lundquist number of the system is defined to be

$$S = \frac{\tau_R}{\tau_A} = \mu_0 \sigma V_A d \mu_0 \eta^{-1} V_A d. \quad (6.8.5)$$

The Lundquist number is a dimensionless number very similar to the magnetic Reynolds number defined earlier by equation (6.3.6), except that U is replaced by V_A . The Lundquist number for hot fusion plasmas, solar coronal plasmas, or magnetotail plasmas can lie in the range $10^8 - 10^{14}$, and so there is a very large separation between τ_A and τ_R in such systems. What is interesting about magnetic reconnection in a resistive plasma is that its

characteristic time scale, τ_r , lies between τ_A and τ_R , i.e., $\tau_A \ll \tau_r \ll \tau_R$. This allows flux annihilation to occur much faster when field lines reconnect than when they diffuse in a high- S plasma. The geometry of magnetic field lines plays a crucial role in determining whether such rapid flux annihilation can occur. In the geometry represented in Figure 6.33, one intuitively expects dramatic changes to occur in a small region $\Delta \ll d$ near $y = 0$. One expects a characteristic reconnection time $\tau_r = \mu_0 \eta^{-1} \Delta^2$ which is much shorter than the resistive diffusion time τ_R , defined by equation (6.8.4). This is borne out by the Sweet–Parker model of magnetic reconnection, discussed in the next section.

6.8.1 The Sweet–Parker model

Parker [1957] and Sweet [1958] developed a simple model of magnetic reconnection that has been a cornerstone of the literature for nearly four decades and has been verified by computer simulations. This model is now known as the Sweet–Parker model. The problem they addressed is the following: consider the field configuration given by equation (6.8.1) in which the component B_x tends to $\pm B_0 \hat{x}$ as $y \rightarrow \pm\infty$. If these oppositely directed fields are convected inward toward each other, what is the speed U_{in} (along y) at which the oppositely directed fields merge and annihilate each other? We assume that the plasma is incompressible. It can then be shown that the field B_T plays no role in the dynamics.

As shown in Figure 6.34, the inward flow U_{in} induces a steady outward flow U_{out} in a narrow channel of width Δ (along x). In steady state, the momentum balance condition along x gives

$$\rho_m U_x \frac{\partial U_x}{\partial x} = -\frac{\partial P}{\partial x}. \quad (6.8.6)$$

Note that variation of the magnetic field along x is neglected. In an incompressible plasma, the density ρ_m is constant, and we can integrate equation (6.8.6) from $x = 0$ to a value of x sufficiently large as to lie just outside the reconnection region. Let this asymptotic distance be $x = L/2$. We then obtain

$$\rho_m \frac{U_{out}^2}{2} = \delta P \quad (6.8.7)$$

where U_{out} is the value of U_x in the asymptotic region, $U_x = 0$ at $x = 0$, and δP is the pressure differential of the fluid between $x = 0$ and $L/2$. This pressure differential is essentially set up by the equilibrium condition (6.6.48), $P + B^2/2\mu_0 = \text{constant}$, across the width of the sheet, with P largest at the neutral line where $B = 0$. Hence, we obtain $\delta P = B_0^2/2\mu_0$ which, substituted in equation (6.8.7), gives

$$U_{out} = \left(\frac{B_0}{\mu_0 \rho_m} \right)^{1/2} = V_A. \quad (6.8.8)$$

By mass continuity of the incompressible fluid in steady state, we obtain

$$U_{in}L = U_{out}\Delta = V_A\Delta. \quad (6.8.9)$$

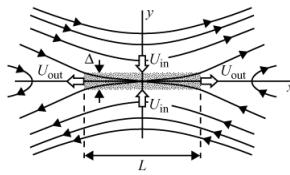


Fig. 6.34 The magnetic field geometry used to evaluate reconnection near an x-type neutral line.

During reconnection, a thin and intense current sheet develops over a narrow width and represents the main site for Ohmic dissipation. In steady state, the resistive dissipation rate in the reconnection layer of width Δ must balance the energy inflow, i.e.,

$$\eta J^2 \Delta = U_{in} \frac{B_0^2}{2\mu_0}, \quad (6.8.10)$$

where

$$J = \frac{B_0}{\mu_0 \Delta}. \quad (6.8.11)$$

From equations (6.8.10) and (6.8.11), we obtain

$$U_{in} = \frac{2\eta}{\mu_0 \Delta}. \quad (6.8.12)$$

Eliminating U_{in} between equations (6.8.9) and (6.8.12), it follows that

$$\Delta = \left(\frac{2\eta L}{V_A \mu_0} \right)^{1/2} = \frac{2L}{S^{1/2}}, \quad (6.8.13)$$

where the Lundquist number S is defined with $2L$ playing the role of the characteristic system size. Substituting expression (6.8.13) for Δ into equation (6.8.12), we obtain

$$U_{in} = \frac{2V_A}{S^{1/2}}. \quad (6.8.14)$$

Note that for $S \gg 1, \Delta \gg 2L$, as anticipated earlier. Hence, the time scale of reconnection is shorter than the time scale of diffusion. Note that this separation of time scales depends crucially on the geometry of the magnetic field lines. If the magnetic field did not point in opposite directions (across the neutral line) but in the same direction, then the parameter Δ in equations (6.8.10) and (6.8.11) would be replaced by L , with the consequence that $U_{in} = 2V_A/S$, which is indistinguishable from simple diffusion. Thus, the reconnection velocity U_{in} is $S^{1/2}$ times larger than the diffusion velocity. Note, however, this large U_{in} is still smaller than the Alfvén speed V_A by a factor of $S^{1/2}$.

6.8.2 Validity of resistive MHD

In deriving the complete set of MHD equations, given in Section 6.1.5, we have explicitly stated three important assumptions. The first is the assumption of a local Maxwellian distribution that leads to an equation of state of the form (6.1.25). This equation of state essentially describes the adiabatic evolution of a plasma and neglects the effects of heat transport, as well as viscous heating. The assumption of a local Maxwellian distribution is valid in the highly collisional regime, i.e., when $L \gg \lambda_{mfp}$, where λ_{mfp} is the collisional mean-free path of the system. The second, and less serious, assumption is the neglect of the displacement current term in Ampère's law (6.1.28) that is valid in a non-relativistic plasma as long as the phase velocity of the waves is much smaller than the speed of light. Finally, the third assumption is that of quasi-neutrality or zero charge density, valid when $L \gg \lambda_D$.

In addition to the three assumptions mentioned above, we had forewarned the reader that the conditions under which the generalized Ohm's law (6.1.23) reduces to the resistive Ohm's law (6.1.13) are quite stringent. Of the three neglected terms on the right-hand side of equation (6.1.23), the neglect of the last term is easiest to justify: it is valid to do so when $m_e \ll m_i$ and $\rho_c \ll L$, where ρ_c is the ion cyclotron radius (??). However, the neglect of the second and third terms, which are called finite Larmor radius terms, is valid only when $\rho_c \ll L$. It turns out that the neglect of these terms is much more delicate than neglect of the electron inertia term. This is especially so in magnetic reconnection problems where the reconnection layer width Δ can often be comparable to ρ_c in higher density plasmas. In such cases, $\rho_c \gtrsim \Delta$ even though $\rho_c \ll L$, which requires that the physics of the reconnection layer be treated by adding the finite Larmor radius terms to the resistive Ohm's law, which now becomes

$$\mathbf{E} + \mathbf{U} \times \mathbf{B} = \eta \mathbf{J} + \frac{\mathbf{J} \times \mathbf{B}}{ne} - \frac{\nabla \cdot \overleftrightarrow{\mathbf{P}}_e}{ne}. \quad (6.8.15)$$

In weakly collisional plasmas, the resistivity η is very low and the Lundquist number S is very high. In such cases it is easy to obtain situations where the parameter Δ , defined by equation (6.8.13), becomes sufficiently small that it is important to replace the resistive Ohm's law by the generalized Ohm's law (6.8.15). In what follows, we will describe qualitatively some of the changes brought about in the reconnection dynamics by the presence of the Hall current term $\mathbf{J} \times \mathbf{B}/ne$. For simplicity, we will consider the case in which the out-of-plane component B_T is zero and neglect the effect of the electron pressure tensor $\overleftrightarrow{\mathbf{P}}_e$. Using the equation $\mathbf{J} = ne(\mathbf{U}_i - \mathbf{U}_e)$ and the approximation $\mathbf{U} \approx \mathbf{U}_i$, we can combine the term $\mathbf{U} \times \mathbf{B}$ on the left and the Hall current term on the right to rewrite equation (6.8.15) in the form

$$\mathbf{E} + \mathbf{U}_e \times \mathbf{B} \approx \eta \mathbf{J}. \quad (6.8.16)$$

In magnetic reconnection, the effect of the resistivity η is felt most strongly in a small

region Δ near the neutral surface $y = 0$, also known as the diffusion region. Away from the diffusion region, the effect of η can be neglected whereby equation (6.8.16) reduces to

$$\mathbf{E} + \mathbf{U}_e \times \mathbf{B} = 0. \quad (6.8.17)$$

The above equation implies that in the presence of the Hall current, the electron fluid, which moves with the velocity \mathbf{U}_e , is frozen to the magnetic field, but the ion fluid, which moves with the velocity \mathbf{U}_i , is not. This is one of the principal signatures of Hall MHD reconnection; the ions and electrons move as separate fluids.

Because the electron fluid is frozen to the magnetic field \mathbf{B} , it follows the magnetic field all the way to the diffusion region and ends up carrying most of the current in this region. The ions, which are not tied to the magnetic field, flow in a channel much broader than the electron channel. Computer simulations show that, under such conditions, the merging flow U_{in} , which should be identified with the ion merging flow, can exceed the Sweet-Parker result U_{in} significantly and depends rather weakly on the Lundquist number S . For more details, the reader is referred to the monograph by Biskamp (2000).

Problems

- 6.1. Show for a steady-state flow ($\partial/\partial t = 0$) that the momentum equation with a gravitational force, $\mathbf{F}_g = \rho_m \mathbf{g}$, an isotropic pressure, and no electromagnetic forces ($\mathbf{E} = 0$ and $\mathbf{B} = 0$)

$$\rho_m \left[\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{U} \cdot \nabla) \mathbf{U} \right] = -\nabla P + \rho_m \mathbf{g}$$

can be integrated once along a streamline to obtain Bernoulli's equation

$$\frac{1}{2} \rho_m U^2 + P + \rho_m g z \simeq \text{constant.}$$

Assume that the gravitational force term can be derived from a potential, $\rho_m \mathbf{g} = -\nabla \Phi_g$, where $\Phi_g = \rho_m g z$ and g is constant.

- 6.2. If the diffusion coefficient κ is a constant, show that the scalar diffusion equation

$$\frac{\partial F}{\partial t} = \kappa \nabla^2 F$$

has a spherically symmetric solution given by

$$F = \left(\frac{1}{4\pi\kappa t} \right)^{3/2} \exp \left[-\frac{r^2}{4\kappa t} \right].$$

Make a sketch of $F(r)$ at two times t_1 and t_2 .

- 6.3. Assuming pressure balance, $B^2/(2\mu_0) + P = \text{constant}$, show that the following function

$$\mathbf{B} = B_0 \operatorname{erf} \left\{ \frac{1}{2} \left(\frac{\mu_0 \sigma}{t} \right)^{1/2} y \right\} \hat{\mathbf{x}}$$

is a solution of the static ($\mathbf{U} = 0$) time dependent equation

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B},$$

where

$$\text{erf}(y) = \frac{1}{\sqrt{\pi}} \int_0^y e^{-v^2} dv.$$

(a) Make a plot of $B_x(y, t)$ as a function of y for a series of times t_1, t_2 , and t_3 .

(b) Compute the current $J_z(y, t)$ and plot as a function of y for a series of times t_1, t_2 , and t_3 .

(c) Make a plot of the pressure function, $P(y)$, that would be required to achieve a force balanced equilibrium for this magnetic field.

- 6.4. If the magnetic flux is frozen in the plasma and if the 1st and 2nd adiabatic invariants are conserved, show that

$$\frac{P_\perp}{\rho_m B} = \text{constant}, \quad \text{and} \quad \frac{P_\parallel P_\perp^2}{\rho_m^5} = \text{constant}.$$

Hint: Note that $P_\perp \sim \rho_m w_\perp$ and that $P_\parallel \sim \rho_m w_\parallel$. The 1st invariant implies that $w_\perp/B = \text{constant}$ and the 2nd invariant implies that $w_\parallel L^2 = \text{constant}$.

- 6.5. If the magnetic field in the solar wind is “frozen” in the plasma, show that the magnetic field near the solar equatorial plane takes on the shape of an Archimedian spiral

$$\phi = \frac{\omega_s}{V_{sw}} \rho + \phi_0,$$

where ϕ is an azimuthal angle viewed from the polar axis, ω_s is the angular rotational velocity of the Sun, ρ is the distance from the Sun, and V_{sw} is the solar wind velocity.

- 6.6. To find solutions of a force-free MHD equilibrium, we frequently look for solutions of the form $\nabla \times \mathbf{B} = \alpha \mathbf{B}$, where α is a constant.

(a) In rectangular coordinates, show that $\mathbf{B} = (0, B_0 \sin \alpha x, B_0 \cos \alpha x)$ is a solution.

(b) In cylindrical coordinates, show that $\mathbf{B} = [B_\rho, B_\phi, B_z] = [0, B_0 J_1(\alpha \rho), B_0 J_0(\alpha \rho)]$ is a solution where J_0 and J_1 are the zeroth- and first-order Bessel functions. Hint: You will need to research the properties of Bessel functions to do this problem. Make a sketch of the magnetic field lines in the region $0 < \alpha \rho < 2.4$ (i.e., inside the first zero of $J_0(\alpha \rho)$).

- 6.7. Show that a time dependent solution of the force-free ($\nabla \times \mathbf{B} = \alpha \mathbf{B}$) static equilibrium equation

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\mu_0 \sigma} \nabla^2 \mathbf{B}$$

is given by

$$\mathbf{B} = \mathbf{B}_0 e^{-t/\tau}$$

where \mathbf{B}_0 is the solution of the vector Helmholtz equation

$$\nabla^2 \mathbf{B}_0 + \alpha^2 \mathbf{B}_0 = 0$$

and $\tau = (\mu_0\sigma)/\alpha^2$, with $\alpha = \text{constant}$.

Note: Solutions of the vector Helmholtz equation can be found in many electricity and magnetism books.

- 6.8. Solutions for force-free MHD equilibria can also be found for which α is not a constant.
 (a) In cylindrical coordinates, show that

$$\mathbf{B} = [B_\rho, B_\phi, B_z] = \left[0, \frac{B_0 k \rho}{1 + k^2 \rho^2}, \frac{B_0}{1 + k^2 \rho^2} \right]$$

is a solution where k is a constant. In this case show that

$$J_z = \frac{2k B_0 / \mu_0}{(1 + k^2 \rho^2)^2}.$$

(b) Make a sketch of B_ϕ , B_z , and J_z as a function of ρ .

- 6.9. For the “Bennett pinch” type of pressure balanced equilibrium, the solution for $n(\rho)$ depends on the conductivity relation. Assuming the thermal conductivity is so large that the temperature is independent of radius, analyze the following cases.
 (a) Lorentz type conductivity. Using $J_z = \frac{e^2 n}{mv} E_z$ show that $n(\rho) = n_0/[1 + (\rho/R)^2]^2$, where R is a constant.
 (b) Fully ionized gas. Using $J_z = \sigma E$, where σ is a constant, show that $n(\rho) = n_0(1 - \rho^2/R^2)$.

- 6.10. In the derivation of small amplitude waves in an MHD fluid, we obtained the equation

$$\omega^2 \mathbf{U} = \frac{1}{\mu_0 \rho_{m0}} \{ \mathbf{k} \times (\mathbf{k} \times [\mathbf{U} \times \mathbf{B}_0]) \} \times \mathbf{B}_0 + V_s^2 \mathbf{k} (\mathbf{k} \cdot \mathbf{U}).$$

- (a) Show that this equation can be represented by the following homogeneous equation:

$$\begin{bmatrix} v_p^2 - V_A^2 - V_S^2 \sin^2 \theta & 0 & -V_S^2 \sin \theta \cos \theta \\ 0 & v_p^2 - V_A^2 \cos^2 \theta & 0 \\ -V_S^2 \sin \theta \cos \theta & 0 & v_p^2 - V_S^2 \cos^2 \theta \end{bmatrix} \begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix} = 0,$$

where $V_A^2 = B_0^2/\mu_0 \rho_{m0}$ and $v_p = \omega/k$.

- (b) Show that the determinant of the above matrix is

$$D(k, \omega) = (v_p^2 - V_A^2 \cos^2 \theta)[v_p^4 - v_p^2(V_A^2 + V_S^2) + V_A^2 V_S^2 \cos^2 \theta] = 0.$$

- 6.11. If we add viscous effects to the MHD momentum equation by including a term $\nu \nabla^2 \mathbf{U}$, where ν is the dynamic viscosity,

$$\rho_m \frac{d\mathbf{U}}{dt} = \mathbf{J} \times \mathbf{B} - \nabla P + v \nabla^2 \mathbf{U}$$

and finite conductivity effects via the term $(1/\mu_0\sigma)\nabla^2 B$ in

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B}) + \frac{1}{\mu_0\sigma} \nabla^2 \mathbf{B},$$

show that the dispersion relation for Alfvén waves can be obtained by multiplying both ω^2 and V_s^2 by a factor $(1 + ik^2/\mu_0\sigma\omega)$ and ω^2 by a factor $(1 + ivk^2/\rho_{m0}\omega)$.

- 6.12. If the finite conductivity and viscous corrections are small (i.e., $\sigma \rightarrow \infty$ and $v \rightarrow 0$), use the results from problem 6.11 to show that, for parallel ($\theta = 0$) propagation, the dispersion relation for the transverse Alfvén wave is

$$k \simeq \frac{\omega}{V_A} + i \frac{\omega^2}{2V_A^3} \left(\frac{1}{\mu_0\sigma} + \frac{v}{\rho_{m0}} \right).$$

Note that the viscous and finite conductivity effects increase rapidly with increasing frequency. Give a simple explanation of this frequency dependence.

- 6.13. Show that the linear force operator

$$\begin{aligned} \mathbf{F}(\xi) = & \frac{1}{\mu_0} [(\nabla \times \{\nabla \times (\xi \times \mathbf{B}_0)\}) \times \mathbf{B}_0 + (\nabla \times \mathbf{B}_0) \times \{\nabla \times (\xi \times \mathbf{B}_0)\}] \\ & + \nabla[\xi \cdot \nabla P_0 + \gamma P_0(\nabla \cdot \xi)], \end{aligned}$$

is self-adjoint, i.e., if there are two eigenfunctions η_1 and η_2 , satisfying standard boundary conditions, then

$$\int_v \eta_2 \cdot \mathbf{F}(\eta_1) d^3x = \int_v \eta_1 \cdot \mathbf{F}(\eta_2) d^3x.$$

References

- Alexandroff, P., and Hopf, H. 1935. *Topologie*. Springer Verlag:
- Alfvén, H. 1942. Existence of electromagnetic-hydrodynamic waves. *Nature* **150**, 404.
- Babcock, H. W. 1961. The topology of the Sun's magnetic field and the 22-year solar cycle. *Astrophys. J.* **133**, 572–587.
- Basu, S., and Kelley, M. C. 1979. A review of recent observations of equatorial scintillations and their relationship to current theories of F region irregularity generation. *Radio Sci.* **14**, 471–485.
- Batchelor, G. K. 1967. *An Introduction to Fluid Mechanics*. Cambridge University Press: Cambridge, p. 74.
- Belcher, J. W., Davis, L., and Smith, E. J. 1969. Large-amplitude Alfvén waves in the interplanetary medium: Mariner 5. *J. Geophys. Res.* **74**, 2302–2308.
- Bennett, W. H. 1934. Magnetically self-focussing streams. *Phys. Rev.* **45**, 890–897.
- Bigg, E. K. 1964. Influence of the satellite Io on Jupiter's decametric emission. *Nature* **203**, 1008–1010.
- Connerney, J. E. P., Baron, R., Satoh, T., and Owen, T. 1993. Images of excited H₃⁺ at the foot of the Io flux tube in Jupiter's atmosphere. *Science* **262**, 1035–1038.
- Cowling, T. G. 1957. *Magnetohydrodynamics*. Interscience Publishers: New York, p. 2.
- Davis, L. 1972. The configuration of the interplanetary magnetic field. In *Solar Terrestrial Physics*, Part II, ed. E. R. Dyer. Reidel: Dordrecht, pp. 32–48.
- Elsasser, W. M. 1956. Hydromagnetic dynamo theory. *Rev. Mod. Phys.* **28**, 135–163.
- Fetter, A. L., and Walecka, J. D. 1980. *Theoretical Mechanics of Particles and Continua*. McGraw-Hill: New York, p. 345.
- Freidberg, J. P. 1987. *Ideal Magnetohydrodynamics*. Plenum Press: New York.
- Goldreich, P., and Lynden-Bell, D. 1969. Io, a Jovian unipolar inductor. *Astrophys. J.* **156**, 59–78.
- Kruskal, M. D., and Schwarzschild, M. 1954. Some instabilities of a completely ionized plasma. *Proc. R. Soc., London, Ser. A* **223**, 348–360.
- Lundquist, S. 1949. Experimental demonstration of magneto-hydrodynamic waves. *Nature* **164**, 145–146.
- Ness, N. F., Acuna, M., Lepping, R. P., Burlaga, L. F., and Behannon, K. W. 1979. Magnetic field studies at Jupiter by Voyager 1: Preliminary results. *Science* **204**, 982–987.
- Parker, E. N. 1957. Sweet's mechanism for merging magnetic fields in conducting fluids. *J. Geophys. Res.* **62**, 509–520.
- Priest, E., and Forbes, T. 2000. *Magnetic Reconnection: MHD Theory and Applications*. Cambridge University Press: Cambridge, p. 261.

- Shafranov, V. D. 1966. Plasma equilibrium in a magnetic field. In *Reviews of Plasma Physics*, Vol. 2, ed. M. A. Leontovich. Consultants Bureau: New York, p. 103.
- Solov'ev, L. S. 1968. The theory of hydrodynamic stability of toroidal plasma configurations. *Sov. Phys. JETP* **26**, 400–407.
- Sweet, P. A. 1958. The production of high energy particles in solar flares, *Nuovo Cimento Suppl. Ser. X* **8**, 188–196.
- Taylor, J. B. 1974. Relaxation of toroidal plasma and generation of reversed magnetic fields. *Phys. Rev. Lett.* **33**, 1139–1141.
- Taylor, J. B. 1986. Relaxation and magnetic reconnection in plasmas. *Rev. Mod. Phys.* **53**, 741–763.
- Vasyliunas, V. M. 1975. Theoretical models of magnetic field line merging, 1. *Rev. Geophys. Space Phys.* **13** (1), 303–336.
- Wilcox, J. M., Boley, F. I., and DeSilva, A. W. 1960. Experimental study of Alfvén wave propagation. *Phys. Fluids* **3**, 15–19.

Further reading

- Biskamp, D. *Magnetic Reconnection in Plasmas*. Cambridge University Press: Cambridge, 2000.
- Chandrasekhar, S. *Hydrodynamic and Hydromagnetic Stability*. Oxford University Press: Oxford, 1961 (reprinted by Dover Publications in 1981).
- Cowling, T. G. *Magnetohydrodynamics*. Wiley Interscience: New York, 1957.
- Freidberg, J. P. *Ideal Magnetohydrodynamics*. Plenum Press: New York, 1987.
- Kivelson, M. G., and Russell, C. T. *Introduction to Space Physics*. Cambridge University Press: Cambridge, 1995, Chapter 2.

Most fluids, including plasmas, have a remarkable tendency to form discontinuities. The formation of discontinuities can usually be traced to a non-linear effect called wave steepening. For example, consider the propagation of a sound wave in an ordinary gas. As we discussed earlier, the propagation speed of a sound wave is given by $V_S^2 = dP/d\rho_m$. For an adiabatic equation of state, $P/\rho_m^\gamma = \text{constant}$, it is easy to show that the propagation speed increases as the pressure increases (V_S is proportional to P^α , where $\alpha = (\gamma - 1)/2\gamma$ is a positive constant). Thus, if a pressure pulse is introduced into the gas, the trailing edge of the pulse, which has a higher pressure, tends to catch up with the leading edge, as shown in Figure 7.1. This amplitude dependence causes the pulse to steepen, until it becomes a nearly discontinuous change in pressure. Discontinuities of this type are called shock waves. Shock waves propagate at a speed that is intermediate between the sound speeds upstream and downstream of the shock. The thickness of a shock is ultimately controlled by some microscopic length scale. In an ordinary gas, the relevant length scale is the mean-free path. In a collisionless plasma, which has an infinite mean-free path, other length scales, such as the ion cyclotron radius, play a crucial role in determining the thickness.

To investigate the types of discontinuities that can exist in an ideal MHD fluid, we start by assuming that a planar discontinuity of zero thickness separates two otherwise uniform fluids. Conservation of mass, momentum, and energy, and the boundary conditions for the electric and magnetic fields are then used to put constraints on the fluid parameters and fields across the discontinuity. These constraints are called the jump conditions. The derivation of the MHD jump conditions is discussed below.

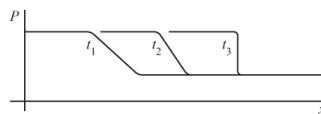


Fig. 7.1 Pulse shapes at a sequence of times t_1 , t_2 , and t_3 that illustrate wave steepening.

7.1 The MHD jump conditions

To derive the jump conditions for an MHD fluid, we start with the laws of conservation of mass, momentum and energy, given by equations (??), (??), and (??), which after some minor rearrangements can be written in the following forms.

1. Conservation of mass

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \mathbf{U}) = 0. \quad (7.1.1)$$

2. Conservation of momentum

$$\frac{\partial}{\partial t}(\rho_m \mathbf{U}) + \nabla \cdot \left[\rho_m \mathbf{U} \mathbf{U} + \left(P + \frac{B^2}{2\mu_0} \right) \hat{\mathbf{1}} - \frac{\mathbf{B} \mathbf{B}}{\mu_0} \right] = 0. \quad (7.1.2)$$

3. Conservation of energy

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho_m U^2 + \xi + \frac{B^2}{2\mu_0} \right) + \nabla \cdot \left(\frac{1}{2} \rho_m U^2 \mathbf{U} + h \mathbf{U} + \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right) = 0, \quad (7.1.3)$$

where ξ is the internal energy, and h is the enthalpy. Since we want to relate the conditions on opposite sides of the discontinuity, consider a pill-box shaped volume V containing the discontinuity, as shown in Figure 7.2. The discontinuity surface divides the volume V into two volumes V_1 and V_2 , each with outer surfaces A_1 and A_2 as shown. The normal to the discontinuity surface, $\hat{\mathbf{n}}$, is assumed to be directed from side 1 to side 2, and the cross-sectional area of the discontinuity surface is A' .

To derive the first jump condition we integrate the mass conservation equation over the volume V

$$\int_V \left[\frac{\partial \rho_m}{\partial t} + \nabla \cdot (\rho_m \mathbf{U}) \right] d^3x = 0. \quad (7.1.4)$$

Using the divergence theorem, the integral involving the divergence can be converted to surface integrals over A_1 and A_2 , and the integral involving the time derivative can be split into two volume integrals over V_1 and V_2 ,

$$\int_{V_1} \frac{\partial \rho_m}{\partial t} d^3x + \int_{V_2} \frac{\partial \rho_m}{\partial t} d^3x + \int_{A_1} (\rho_m \mathbf{U}) \cdot d\mathbf{A} + \int_{A_2} (\rho_m \mathbf{U}) \cdot d\mathbf{A} = 0. \quad (7.1.5)$$

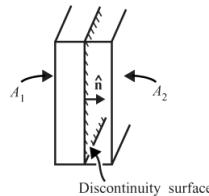


Fig. 7.2 The pill-box shaped surface used to derive the MHD jump conditions.

Next, let the volumes V_1 and V_2 shrink to zero in such a way that the area A' remains constant. The first and second terms (the volume integrals) then go to zero. Since $A_1 \rightarrow A'$ and $A_2 \rightarrow A'$ the third and fourth terms (the surface integrals) can be rewritten as

$$\int_{A'} (\rho_{m1} \mathbf{U}_1 \cdot (-\hat{\mathbf{n}})) dA' + \int_{A'} (\rho_{m2} \mathbf{U}_2 \cdot \hat{\mathbf{n}}) dA' = 0, \quad (7.1.6)$$

or

$$\int_{A'} (\rho_{m1} \mathbf{U}_1 \cdot \hat{\mathbf{n}}) dA' = \int_{A'} (\rho_{m2} \mathbf{U}_2 \cdot \hat{\mathbf{n}}) dA', \quad (7.1.7)$$

where $\hat{\mathbf{n}}$ is a unit vector normal to the surface of the discontinuity. Since the above result is true for arbitrary A' , the integrands must be equal, which gives the first jump condition, *conservation of mass flux*,

$$\rho_{m1} \mathbf{U}_1 \cdot \hat{\mathbf{n}} = \rho_{m2} \mathbf{U}_2 \cdot \hat{\mathbf{n}}. \quad (7.1.8)$$

Jump conditions are often represented using the symbol $\{\}$, which denotes the difference of the enclosed quantity between the two sides of the discontinuity. Using the jump symbol the above equation can be written

$$\{\rho_m \mathbf{U} \cdot \hat{\mathbf{n}}\} = 0. \quad (7.1.9)$$

Note that the above jump condition can be obtained by simply replacing the divergence operator, $\nabla \cdot ()$, in the mass continuity equation with the unit normal, and then setting the jump in this quantity equal to zero, i.e., $\{\hat{\mathbf{n}} \cdot ()\} = 0$. Following this procedure, it is straightforward to obtain the second jump condition, which is *conservation of momentum flux*

$$\left\{ \rho_m \mathbf{U}(\mathbf{U} \cdot \hat{\mathbf{n}}) + \left(P + \frac{B^2}{2\mu_0} \right) \hat{\mathbf{n}} - \frac{\mathbf{B}}{\mu_0} (\mathbf{B} \cdot \hat{\mathbf{n}}) \right\} = 0, \quad (7.1.10)$$

and the third jump condition, which is *conservation of energy flux*

$$\left\{ \left(\frac{1}{2} \rho_m U^2 + h \right) (\mathbf{U} \cdot \hat{\mathbf{n}}) + \frac{1}{\mu_0} (\mathbf{E} \times \mathbf{B}) \cdot \hat{\mathbf{n}} \right\} = 0. \quad (7.1.11)$$

In addition to the above jump conditions, Maxwell's equations require that the tangential component of the electric field and the normal component of the magnetic field must be continuous across the discontinuity,

$$\{\mathbf{E}_t\} = 0 \quad (7.1.12)$$

and

$$\{\mathbf{B} \cdot \hat{\mathbf{n}}\} = 0. \quad (7.1.13)$$

The above set of five jump conditions is sufficient for determining the possible discontinuities that can exist in an MHD fluid.

The jump conditions can be put into a more useful form by introducing the subscripts n and t for the normal and tangential components relative to the discontinuity surface and eliminating the electric field using $\mathbf{E} = -\mathbf{U} \times \mathbf{B}$. After some minor manipulations it is easy to show that the jump conditions can be written

$$1. \quad \{\rho_m U_n\} = 0 \quad (7.1.14)$$

$$2a. \quad \left\{ \rho_m U_n^2 + P + \frac{B_t^2}{2\mu_0} \right\} = 0 \quad (7.1.15)$$

$$2b. \quad \left\{ \rho_m U_n \mathbf{U}_t - B_n \frac{\mathbf{B}_t}{\mu_0} \right\} = 0 \quad (7.1.16)$$

$$3. \quad \left\{ \left(\frac{1}{2} \rho_m U^2 + h + \frac{B^2}{\mu_0} \right) U_n - (\mathbf{U} \cdot \mathbf{B}) \frac{B_n}{\mu_0} \right\} = 0 \quad (7.1.17)$$

$$4. \quad \{\mathbf{U}_n \times \mathbf{B}_t + \mathbf{U}_t \times \mathbf{B}_n\} = 0 \quad (7.1.18)$$

$$5. \quad \{B_n\} = 0. \quad (7.1.19)$$

Before proceeding further, it is helpful to note certain simple algebraic rules concerning the jump symbol, $\{\}$. First, as can be easily verified, the jump symbol is distributive, i.e., $\{a + b\} = \{a\} + \{b\}$. Thus, $\{B^2\} = \{B_n^2 + B_t^2\} = \{B_n^2\} + \{B_t^2\}$. Since jump condition 5 above implies that $\{B_n^2\} = 0$, it follows that $\{B^2\} = \{B_t^2\}$. This result was used in the derivation of jump condition 2b above. Second, whenever some quantity, such as $(\rho_m U_n)$ is known to be conserved, i.e., $\{\rho_m U_n\} = 0$, this quantity can be factored out of any term in which it occurs. Thus, in jump condition 2b, we can write $\{\rho_m U_n \mathbf{U}_t\} = (\rho_m U_n) \{\mathbf{U}_t\}$. Since the conserved quantity is the same on both sides of the discontinuity, it is not necessary to specify on which side of the discontinuity the quantity is to be evaluated. However, if the conserved quantity consists of a product of two or more terms, these terms must all be evaluated on the same side of the discontinuity.

7.2 Classification of discontinuities

To analyze the types of discontinuities that can occur in an MHD fluid, it is convenient to categorize them according to whether U_n and $\{\rho_m\}$ are zero or non-zero. This classification scheme was first introduced by Landau and Lifshitz [1960]. Of the four possible combinations, the case $U_n = 0$ and $\{\rho_m\} = 0$ is trivial, since no change occurs across the discontinuity. The remaining three cases (see Table 7.1) are called a contact discontinuity ($U_n = 0, \{\rho_m\} \neq 0$), a rotational discontinuity ($U_n \neq 0, \{\rho_m\} = 0$), and a shock wave ($U_n \neq 0, \{\rho_m\} \neq 0$). Each of these cases will be analyzed separately. Sometimes it is also useful to consider further subcategories depending on whether B_n is zero or non-zero.

Table 7.1 Classification of MHD discontinuities

$U_n = 0$	$U_n \neq 0$
$\{\rho_m\} = 0$	trivial
$\{\rho_m\} \neq 0$	contact discontinuity

rotational discontinuity
shock wave

7.2.1 Contact discontinuities

Contact discontinuities are characterized by no fluid flow across the discontinuity, $U_n = 0$, and a non-zero density jump, $\{\rho_m\} \neq 0$. If $B_n \neq 0$ it is easily verified from the jump conditions that $\{U_t\} = 0$, $\{B\} = 0$, and $\{P\} = 0$. Thus, only the density can change across the discontinuity. Because the pressure cannot change the density change must be compensated by a corresponding change in temperature such that $n_1\kappa T_1 = n_2\kappa T_2$. This type of discontinuity is called a *contact discontinuity*. Contact discontinuities are often encountered in explosive processes, such as chemical explosions, where a new source of hot gas is suddenly injected into the system. Although contact discontinuities are common in ordinary gases, such discontinuities (with $B_n \neq 0$) are rarely observed in plasmas, since the plasma can readily diffuse along the magnetic field between the two regions.

A special case of a contact discontinuity occurs when $B_n = 0$. If $B_n = 0$, it is easy to show from the jump conditions that $\{U_t\} \neq 0$, $\{B_t\} \neq 0$, and

$$\left\{ P + \frac{B^2}{2\mu_0} \right\} = 0. \quad (7.2.1)$$

This special case of a contact discontinuity is called a *tangential discontinuity*. For a tangential discontinuity, the fluid velocity and magnetic field are parallel to the surface of the discontinuity but change in magnitude or direction. The above equation shows that the sum of the plasma and magnetic field pressures must be constant across the discontinuity.

Since plasmas cannot easily diffuse across magnetic field lines, tangential discontinuities are commonly observed in plasmas. A good example occurs in association with planetary magnetospheres. Planetary magnetospheres typically have a well-defined boundary between the solar wind and the planetary magnetic field. This boundary is called the magnetopause [Parks, 1991]. If the reconnection rate between the planetary and solar wind magnetic fields is small, then the normal components of the flow velocity and the magnetic field at the magnetopause are also small ($U_n \approx 0$ and $B_n \approx 0$). The magnetosphere is then said to be closed, since the solar wind plasma and magnetic field do not penetrate into the magnetosphere, see Figure 7.3. Under these conditions the magnetopause is to a good approximation a tangential discontinuity. The plasma and magnetic field pressures on the two sides of the magnetopause must then satisfy the pressure balance condition given by equation (7.2.1). Tangential discontinuities are also sometimes observed separating regions with different pressures and magnetic fields in the solar wind [Smith, 1973].

7.2.2 Rotational discontinuities

Rotational discontinuities are characterized by a fluid flow across the discontinuity, $U_n \neq 0$, and no change in the density, $\{\rho_m\} = 0$. Since the density is constant it immediately follows that the enthalpy and pressure are unchanged across the discontinuity, $\{h\} = 0$ and $\{P\} = 0$. From jump conditions 1 and 2a, it is also easily shown that $\{U_n\} = 0$ and $\{B_t^2\} = 0$.

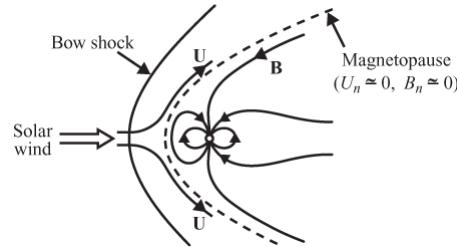


Fig. 7.3

The interaction of the solar wind with a magnetized planet produces two types of discontinuities, the *bow shock* where the supersonic solar wind becomes subsonic, and the *magnetopause* where the planetary magnetic field pressure balances the solar wind pressure.

From jump conditions 2b and 4, one can see that $\{\mathbf{U}_t\}$ and $\{\mathbf{B}_t\}$ must simultaneously satisfy

$$(\rho_m U_n) \{\mathbf{U}_t\} = \frac{B_n}{\mu_0} \{\mathbf{B}_t\} \quad (7.2.2)$$

and

$$\mathbf{U}_n \times \{\mathbf{B}_t\} + \{\mathbf{U}_t\} \times \mathbf{B}_n = 0. \quad (7.2.3)$$

Eliminating $\{\mathbf{U}_t\}$ from the above equations, we obtain

$$\mathbf{U}_n \times \{\mathbf{B}_t\} = B_n \frac{\mathbf{B}_n \times \{\mathbf{B}_t\}}{\mu_0 (\rho_m U_n)}, \quad (7.2.4)$$

which is satisfied identically if

$$U_n^2 = \frac{B_n^2}{\mu_0 \rho_m}. \quad (7.2.5)$$

This type of discontinuity is called a *rotational discontinuity* since the magnetic field \mathbf{B}_t remains constant in magnitude but rotates in the plane of the discontinuity. The field geometry viewed normal to the discontinuity is as shown in Figure 7.4. The speed at which the discontinuity propagates relative to the rest frame of the fluid is $U_n = V_A \cos \theta$, where θ is the angle between the magnetic field and the normal, and $V_A = B / \sqrt{\mu_0 \rho_m}$ is the Alfvén speed.

It is easily recognized that for small amplitudes a rotational discontinuity is simply a

transverse Alfvén wave. In the linearized treatment \mathbf{U}_1 and \mathbf{B}_1 correspond to $\{\mathbf{U}_t\}$ and $\{\mathbf{B}_t\}$, the wave number \mathbf{k} is parallel to the normal, $B_0 \cos \theta$ corresponds to B_n , and $V_A \cos \theta$ corresponds to U_n .

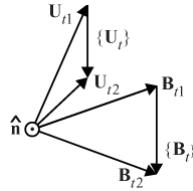


Fig. 7.4 The field geometry of a rotational discontinuity as viewed normal to the plane of the discontinuity.

If the reconnection rate between the solar wind magnetic field and the planetary magnetic field is substantial, as tends to occur when the solar wind magnetic field is directed opposite to the planetary magnetic field, then the solar wind can penetrate significantly into the magnetosphere ($U_n \neq 0$ and $B_n \neq 0$). The magnetosphere is then said to be open. Under these conditions, the magnetopause becomes a rotational discontinuity [Sonnerup *et al.*, 1981]. Rotational discontinuities are also frequently observed in the solar wind [Smith, 1973].

7.3 Shock waves

Shock waves are characterized by a fluid flow across the discontinuity, $U_n \neq 0$, and a non-zero density jump, $\{\rho_m\} \neq 0$. Of the three basic types of discontinuities, shock waves are by far the most difficult to analyze, since all of the jump conditions must be used in the analysis. To simplify the analysis, it is useful to make a specific choice for the coordinate system, called the de Hoffmann–Teller frame of reference, after de Hoffmann and Teller [1950], who first pointed out the advantages of using this coordinate system.

7.3.1 The de Hoffmann–Teller frame

In the de Hoffmann–Teller frame of reference, the fluid velocity \mathbf{U} and the magnetic field \mathbf{B} are parallel both upstream and downstream of the shock, and lie in a plane perpendicular to the discontinuity surface. To prove that such a coordinate system can be found, start with jump conditions 2b and 4. After recognizing that $(\rho_m U_n)$ and B_n are conserved across the discontinuity, these two equations can be written

$$(\rho_m U_n) \{\mathbf{U}_t\} - B_n \left\{ \frac{\mathbf{B}_t}{\mu_0} \right\} = 0, \quad (7.3.1)$$

and

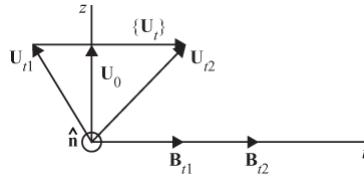


Fig. 7.5 The field geometry of a shock wave viewed normal to the plane of the discontinuity. The normal, $\hat{\mathbf{n}}$, is directed upward out of the paper.

$$(\rho_m U_n) \left\{ \frac{\mathbf{B}_t}{\rho_m} \right\} - B_n \{ \mathbf{U}_t \} = 0. \quad (7.3.2)$$

From the above two equations, it is clear that the vectors $\{\mathbf{B}_t/\rho_m\}$, $\{\mathbf{U}_t\}$, and $\{\mathbf{B}_t\}$ are all parallel. Solving the second equation for $\{\mathbf{U}_t\}$ and substituting it into the first equation then gives the equation

$$(\rho_m U_n)^2 \left\{ \frac{\mathbf{B}_t}{\rho_m} \right\} - B_n^2 \left\{ \frac{\mathbf{B}_t}{\mu_0} \right\} = 0, \quad (7.3.3)$$

which can be rewritten in the form

$$\mathbf{B}_{t1} \left(\frac{(\rho_m U_n)^2}{\rho_{m1}} - \frac{B_n^2}{\mu_0} \right) = \mathbf{B}_{t2} \left(\frac{(\rho_m U_n)^2}{\rho_{m2}} - \frac{B_n^2}{\mu_0} \right). \quad (7.3.4)$$

If $\rho_{m1} \neq \rho_{m2}$ the above equations show that \mathbf{B}_{t1} and \mathbf{B}_{t2} are parallel. From equation (7.3.2) and the definition of $\{\mathbf{B}_t\}$, it then follows that \mathbf{B}_{t1} , \mathbf{B}_{t2} , $\{\mathbf{B}_t\}$, and $\{\mathbf{U}_t\}$ are all parallel to each other and lie in the plane of the shocks. The field geometry viewed normal to the discontinuity surface is then as shown in Figure 7.5, where for notational convenience we have selected the tangential, t , axis to be parallel to \mathbf{B}_{t1} and \mathbf{B}_{t2} . The z axis completes the right-hand (n, t, z) coordinate system. Unless otherwise specified, side 1 is assumed to be the upstream side (i.e., U_{n1} is positive). From the above geometry, it is clear that the fluid velocities \mathbf{U}_{t1} and \mathbf{U}_{t2} can be made parallel to \mathbf{B}_{t1} and \mathbf{B}_{t2} by transforming to a coordinate system moving with a velocity \mathbf{U}_0 parallel to the z axis. Thus, without loss of generality, a coordinate system can be found in which \mathbf{U} and \mathbf{B} lie in the n, t plane. The field geometry in this coordinate system is shown in Figure 7.6.

It is easily verified that the electric field $\mathbf{E} = -\mathbf{U} \times \mathbf{B}$ in this new coordinate system is in the z direction and is given by

$$E_z = -(U_{n1}B_{t1} - U_{t1}B_n) = -(U_{n2}B_{t2} - U_{t2}B_n). \quad (7.3.5)$$

By performing a second transformation to a coordinate system moving with a constant velocity parallel to the t axis, this electric field can be transformed to zero. The appropriate translation velocity is easily shown to be

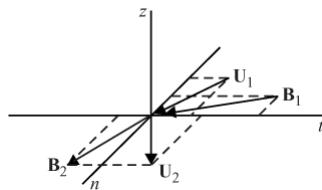


Fig. 7.6 By performing a transformation along the z axis, a coordinate system can be found where the flow velocity \mathbf{U} , and the magnetic field \mathbf{B} , upstream and downstream of the shock, are coplanar.

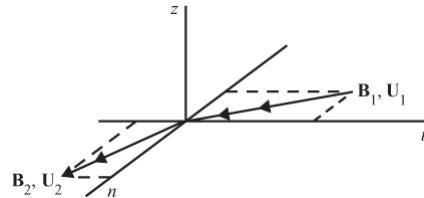


Fig. 7.7 By performing a second transformation, this time along the t axis, a coordinate system can be found where the flow velocity \mathbf{U} and the magnetic field \mathbf{B} are parallel both upstream and downstream of the shock. This coordinate system is called the de Hoffmann–Teller frame.

$$U_{t0} = U_{t1} - U_{n1} \left(\frac{B_{t1}}{B_n} \right). \quad (7.3.6)$$

This is the de Hoffmann–Teller frame. Since $\mathbf{E} = -\mathbf{U} \times \mathbf{B} = 0$, it follows that in this frame of reference \mathbf{U} and \mathbf{B} are parallel. The field geometry is then as shown in Figure 7.7. Thus, by performing two coordinate transformations, we have reduced an inherently three-dimensional problem to two dimensions (i.e., the n, t plane), thereby greatly simplifying the analysis.

7.3.2 The shock propagation speed

The next step in the analysis is to determine the normal component of upstream flow velocity, U_n , as a function of the upstream parameters, thereby giving the propagation speed of the shock relative to the upstream fluid. At this point it is convenient to introduce the speed of sound, which as before is defined by

$$V_S^2 = \frac{\gamma P}{\rho_m}. \quad (7.3.7)$$

In terms of the shock coordinates (n, t, z) described earlier, the jump conditions given by equations (7.1.14)–(7.1.19) can then be written

$$1. \quad \rho_{m1} U_{n1} = \rho_{m2} U_{n2} \quad (7.3.8)$$

$$2a. \quad \rho_{m1} U_{n1}^2 + \rho_{m1} \frac{V_{S1}^2}{\gamma} + \frac{B_{t1}^2}{2\mu_0} = \rho_{m2} U_{n2}^2 + \rho_{m2} \frac{V_{S2}^2}{\gamma} + \frac{B_{t2}^2}{2\mu_0} \quad (7.3.9)$$

$$2b. \quad \rho_{m1} U_{n1} U_{t1} - B_{n1} \frac{B_{t1}}{\mu_0} = \rho_{m2} U_{n2} U_{t2} - B_{n2} \frac{B_{t2}}{\mu_0} \quad (7.3.10)$$

$$3. \quad \begin{aligned} & \rho_{m1} U_{n1} \left(\frac{V_{S1}^2}{\gamma - 1} + \frac{U_{t1}^2 + U_{n1}^2}{2} \right) + \frac{B_{t1}}{\mu_0} (B_{t1} U_{n1} - B_n U_{t1}) \\ &= \rho_{m1} U_{n2} + \left(\frac{V_{S2}^2}{\gamma - 1} + \frac{U_{t2}^2 + U_{n2}^2}{2} \right) + \frac{B_{t2}}{\mu_0} (B_{t2} U_{n2} - B_n U_{t2}) \end{aligned} \quad (7.3.11)$$

$$4a. \quad U_{n1} B_{t1} - U_{t1} B_{n1} = U_{n2} B_{t2} - U_{t2} B_{n2} \quad (7.3.12)$$

$$4b. \quad U_{n2} B_{t2} - U_{t2} B_{n2} = 0 \quad (\text{i.e., the } E_z = 0 \text{ condition}) \quad (7.3.13)$$

$$5. \quad B_{n1} = B_{n2}, \quad (7.3.14)$$

where we have used $h = \gamma P / (\gamma - 1)$ for the enthalpy. These equations are known as the *Rankine–Hugoniot* equations, after a similar set of well-known equations in fluid mechanics [Currie, 1974].

The procedure used to solve the above system of equations is to specify all of the upstream parameters except \mathbf{U} and then to solve for U_{n1} as a function of the upstream parameters. In the above jump conditions, there are twelve unknowns ($\rho_{m1}, \rho_{m2}, U_{n1}, U_{n2}, U_{t1}, U_{t2}, V_{S1}, V_{S2}, B_{t1}, B_{t2}, B_{n1}$, and B_{n2}) of which four upstream parameters ($\rho_{m1}, V_{S1}, B_{t1}$ and B_{n1}) are specified. The jump conditions provide seven equations for $12 - 4 = 8$ unknown quantities. Since the number of unknowns exceeds the number of equations by one, we must specify one more quantity in order to solve the system of equations. This indeterminacy is due to the fact that no information has been provided about the strength of the shock. Ultimately this information involves the boundary conditions on the fluid flow. To make the problem tractable we introduce a new parameter

$$\delta = \frac{\rho_{m2}}{\rho_{m1}}, \quad (7.3.15)$$

called the shock strength, which specifies the strength of the shock. With this parameter specified, enough equations are now available to determine completely all of the unknown parameters.

Before proceeding, to provide some further simplification, it is useful to introduce two dimensionless quantities

$$M_A = \frac{U_n}{V_{An}} = \frac{U_n \sqrt{\mu_0 \rho_m}}{B_n}, \quad (7.3.16)$$

and

$$M_S = \frac{U_n}{V_S} = U_n \sqrt{\frac{\rho_m}{\gamma P}}. \quad (7.3.17)$$

The quantity M_A , is called the Alfvén Mach number and is the ratio of the normal component of the flow velocity to the normal component of the Alfvén velocity. The quantity M_S is called the sonic Mach number and is the ratio of the normal component of the flow speed to the sound speed. It is also convenient to define an angle θ between the magnetic field and the shock normal. This angle is given by

$$\tan \theta = B_t / B_n. \quad (7.3.18)$$

The next step is to eliminate all quantities on side 2 of the jump conditions except the density ρ_{m2} , which will be combined as a ratio with ρ_{m1} to introduce the shock strength δ .

First, we solve for B_{t2} and U_{t2} from jump conditions 2b and 4b. After substituting B_n for B_{n1} and B_{n2} these two jump conditions become

$$\rho_{m2} U_{n2} U_{t2} - \frac{B_n}{\mu_0} B_{t2} = \rho_{m1} U_{n1} U_{t1} - \frac{B_n B_{t1}}{\mu_0}, \quad (7.3.19)$$

and

$$U_{n2} B_{t2} - B_n U_{t2} = 0. \quad (7.3.20)$$

Using the second equation to eliminate U_{t2} from the first equation and solving for B_{t2} gives

$$B_{t2} = \frac{B_n \left(\rho_{m1} U_{n1} U_{t1} - \frac{B_{t1} B_n}{\mu_0} \right)}{\left(\rho_{m2} U_{n2}^2 - \frac{B_n^2}{\mu_0} \right)}. \quad (7.3.21)$$

Next, use jump conditions 4a and 4b ($B_n U_{t1} = U_{n1} B_{t1}$) to eliminate U_{t1} , which gives

$$B_{t2} = \left[\frac{\left(\rho_{m1} U_{n1}^2 - \frac{B_n^2}{\mu_0} \right)}{\left(\rho_{m2} U_{n2}^2 - \frac{B_n^2}{\mu_0} \right)} \right] B_{t1}. \quad (7.3.22)$$

By dividing the numerator and denominator by B_n^2/μ_0 and noting that $\delta = \rho_{m2}/\rho_{m1}$, the above equation can be written in terms of the Alfvén Mach number as follows

$$B_{t2} = \frac{M_{A1}^2 - 1}{\frac{M_{A1}^2}{\delta} - 1} B_{t1}. \quad (7.3.23)$$

Next, the tangential velocity U_{t2} can be determined by substituting B_{t2} from equation (7.3.21) into equation (7.3.20) to give

$$U_{t2} = U_{n2} \frac{\left(\rho_{m1} U_{n1} U_{t1} - \frac{B_n B_{t1}}{\mu_0} \right)}{\left(\rho_{m2} U_{n2}^2 - \frac{B_n^2}{\mu_0} \right)}, \quad (7.3.24)$$

which can be reorganized in the form

$$U_{t2} = U_{t1} + \frac{\left(\frac{B_n^2 U_{t1}}{\mu_0} - \frac{U_{n2} B_n B_{t1}}{\mu_0} \right)}{\left(\rho_{m2} U_{n2}^2 - \frac{B_n^2}{\mu_0} \right)} = U_{t1} + \left\{ \frac{\left[\frac{U_{t1}}{U_{n1}} - \left(\frac{U_{n2}}{U_{n1}} \right) \left(\frac{B_{t1}}{B_n} \right) \right]}{\left(\frac{M_{A1}^2}{\delta} - 1 \right)} \right\} U_{n1}. \quad (7.3.25)$$

Using jump conditions 1 and 4a and introducing $\tan \theta_1 = B_{t1}/B_n$, the above equation can be written in terms of the Alfvén Mach number as follows

$$U_{t2} = U_{t1} + \frac{(\delta - 1) \tan \theta_1}{(M_{A1}^2 - \delta)} U_{n1}. \quad (7.3.26)$$

Having obtained equations for B_{t2} and U_{t2} in terms of quantities on side 1, these equations can then be used to evaluate various quantities in the jump conditions. In jump condition 2a the quantity

$$\frac{B_{t2}^2 - B_{t1}^2}{2\mu_0} = \left[\left(\frac{M_{A1}^2 - 1}{\frac{M_{A1}^2}{\delta} - 1} \right)^2 - 1 \right] \frac{B_{t1}^2}{2\mu_0} \quad (7.3.27)$$

simplifies to

$$\frac{B_{t2}^2 - B_{t1}^2}{2\mu_0} = \frac{(\delta - 1)[(\delta + 1)M_{A1}^2 - 2\delta]M_{A1}^2}{(M_{A1}^2 - \delta)^2} \left(\frac{B_{t1}^2}{2\mu_0} \right). \quad (7.3.28)$$

In jump condition 3, after factoring out $(\rho_{m1} U_{n1}) = (\rho_{m2} U_{n2})$, the quantity

$$\frac{B_{t2}^2}{\mu_0 \rho_{m2}} - \frac{B_{t1}^2}{\mu_0 \rho_{m1}} = \left[\frac{1}{\delta} \left(\frac{M_{A1}^2 - 1}{\frac{M_{A1}^2}{\delta} - 1} \right)^2 - 1 \right] \frac{B_{t1}^2}{\mu_0 \rho_{m1}} \quad (7.3.29)$$

simplifies to

$$\frac{B_{t2}^2}{\mu_0 \rho_{m2}} - \frac{B_{t1}^2}{\mu_0 \rho_{m1}} = \frac{(\delta - 1)(M_{A1}^4 - \delta)}{(M_{A1}^2 - \delta)^2} \left(\frac{B_{t1}^2}{\mu_0 \rho_{m1}} \right). \quad (7.3.30)$$

Similarly, the quantity

$$\begin{aligned}
& \frac{1}{2} (U_{t2}^2 - U_{t1}^2) \\
&= \frac{1}{2} (U_{t2}^2 - U_{t1}^2) \frac{M_{A1}^2 B_n^2}{U_{n1}^2 \mu_0 \rho_{m1}} = \frac{1}{2} (U_{n2}^2 B_{t2}^2 - U_{n1}^2 B_{t1}^2) \frac{M_{A1}^2}{U_{n1}^2 \mu_0 \rho_{m1}} \\
&= \frac{1}{2} \left(\frac{B_{t2}^2}{\delta^2} - B_{t1}^2 \right) \frac{M_{A1}^2}{\mu_0 \rho_{m1}} = \frac{1}{2} \left[\left(\frac{M_{A1}^2 - 1}{M_{A1}^2 - \delta} \right)^2 - 1 \right] \frac{M_{A1}^2 B_{t1}^2}{\mu_0 \rho_{m1}}
\end{aligned} \tag{7.3.31}$$

simplifies to

$$\frac{1}{2} (U_{t2}^2 - U_{t1}^2) = \frac{(\delta - 1)}{2} \left[\frac{2M_{A1}^2 - (1 + \delta)}{(M_{A1}^2 - \delta)^2} \right] \frac{M_{A1}^2 B_{t1}^2}{\mu_0 \rho_{m1}}, \tag{7.3.32}$$

and the quantity

$$\begin{aligned}
\frac{B_n}{\mu_0 \rho_{m1} U_{n1}} (B_{t2} U_{t2} - B_{t1} U_{t1}) &= \frac{B_{t1}}{\mu_0 \rho_{m1} U_{n1}} (B_{t2} U_{t2} - B_{t1} U_{t1}) \\
&= \frac{B_{t1}}{\mu_0 \rho_{m1}} \left[B_{t2} \left(\frac{U_{t2}}{U_{t1}} \right) - B_{t1} \right] \\
&= \left[\frac{1}{\delta} \left(\frac{M_{A1}^2 - 1}{\frac{M_{A1}^2}{\delta} - 1} \right)^2 - 1 \right] \frac{B_{t1}^2}{\mu_0 \rho_{m1}}
\end{aligned} \tag{7.3.33}$$

simplifies to

$$\frac{B_n}{\mu_0 \rho_{m1} U_{n1}} (B_{t2} U_{t2} - B_{t1} U_{t1}) = \frac{(\delta - 1)(M_{A1}^4 - \delta)}{(M_{A1}^2 - \delta)^2} \frac{B_{t1}^2}{\mu_0 \rho_{m1}}. \tag{7.3.34}$$

These equations are then used to eliminate the corresponding quantities from the jump conditions in which they appear. Substituting equation (7.3.28) into jump condition 2a gives

$$\begin{aligned}
& (\rho_{m2} U_{n2}^2 - \rho_{m1} U_{n1}^2) + \left(\rho_{m2} \frac{V_{S2}^2}{\gamma} - \rho_{m1} \frac{V_{S1}^2}{\gamma} \right) + \left(\frac{B_{t1}^2}{2\mu_0} \right) \\
& \times \frac{M_{A1}^2 (\delta - 1)[(\delta + 1)M_{A1}^2 - 2\delta]}{(M_{A1}^2 - \delta)^2} = 0,
\end{aligned} \tag{7.3.35}$$

which, after dividing by $B_n^2/2\mu_0$ and introducing M_A^2 and M_S^2 from equations (7.3.16) and (7.3.17), simplifies to

$$\begin{aligned}
& M_{A2}^2 - M_{A1}^2 + \frac{1}{\gamma} \left(\frac{M_{A2}^2}{M_{S2}^2} - \frac{M_{A1}^2}{M_{S1}^2} \right) \\
& + \frac{M_{A1}^2 (\delta - 1)[(\delta + 1)M_{A1}^2 - 2\delta]}{2(M_{A1}^2 - \delta)^2} \tan^2 \theta_1 = 0,
\end{aligned} \tag{7.3.36}$$

where $\tan^2 \theta_1$ has been substituted for B_{t1}^2/B_n^2 .

Next, dividing jump condition 3 on both sides by $\rho_{m1} U_{n1} = \rho_{m2} U_{n2}$, we obtain

$$\begin{aligned} & \frac{1}{2}(U_{n2}^2 - U_{n1}^2) + \frac{1}{\gamma - 1}(V_{S2}^2 - V_{S1}^2) + \frac{1}{2}(U_{t2}^2 - U_{t1}^2) \\ & + \left(\frac{B_{t2}^2}{\mu_0 \rho_{m2}} - \frac{B_{t1}^2}{\mu_0 \rho_{m1}} \right) - \frac{B_n}{\mu_0 \rho_{m1} U_{n1}} (B_{t2} U_{t2} - B_{t1} U_{t1}) = 0, \end{aligned} \quad (7.3.37)$$

which, after dividing by $B_n^2/\mu_0 \rho_{m2}$, introducing M_A^2 and M_S^2 from equations (7.3.16) and (7.3.17), and substituting equations (7.3.30), (7.3.32), and (7.3.34), simplifies to

$$\begin{aligned} & \frac{1}{2}(M_{A2}^2 - \delta M_{A1}^2) + \frac{1}{\gamma - 1} \left(\frac{M_{A2}^2}{M_{S2}^2} - \frac{M_{A1}^2}{M_{S1}^2} \right) - \frac{(\delta - 1)}{\gamma - 1} \frac{M_{A1}^2}{M_{S1}^2} \\ & + \frac{(\delta - 1)}{2} \frac{(2M_{A1}^2 - \delta - 1)}{(M_{A1}^2 - \delta)^2} M_{A1}^2 \tan^2 \theta_1 = 0, \end{aligned} \quad (7.3.38)$$

where $\tan^2 \theta_1$ has again been substituted for B_{t1}^2/B_n^2 . Finally, the desired equation for M_{A1} is obtained by eliminating the term $[(M_{A2}^2/M_{S2}^2) - (M_{A1}^2/M_{S1}^2)]$ from equations (7.3.36) and (7.3.38), and eliminating the M_{A2}^2/M_{A1}^2 term from the resulting equation using

$$\frac{M_{A2}^2}{M_{A1}^2} = \frac{U_{n2}^2}{(B_n^2/\mu_0 \rho_{m2})} \frac{(B_n^2/\mu_0 \rho_{m1})}{U_{n1}^2} = \left(\frac{\rho_{m1}}{\rho_{m2}} \right)^2 \left(\frac{\rho_{m2}}{\rho_{m1}} \right) = \frac{1}{\delta}. \quad (7.3.39)$$

The result, after some lengthy but straightforward algebra, is

$$\begin{aligned} & (M_{A1}^2 - \delta)^2 \left[M_{S1}^2 - \frac{2\delta}{\delta + 1 - \gamma(\delta - 1)} \right] \\ & - \delta \tan^2 \theta_1 M_{S1}^2 \left[\frac{2\delta - \gamma(\delta - 1)}{\delta + 1 - \gamma(\delta - 1)} M_{A1}^2 - \delta \right] = 0. \end{aligned} \quad (7.3.40)$$

At this point it is convenient to convert the upstream Alfvén Mach number back to the upstream flow velocity using the relation $M_{A1} = U_{n1}/(V_{A1} \cos \theta_1)$. After making this substitution and simplifying, the above equation becomes

$$\begin{aligned} & (U_{n1}^2 - \delta V_{A1}^2 \cos^2 \theta_1)^2 \left[U_{n1}^2 - \frac{2\delta V_{S1}^2}{\delta + 1 - \gamma(\delta - 1)} \right] \\ & - \delta \sin^2 \theta_1 U_{n1}^2 V_{A1}^2 \left[\frac{2\delta - \gamma(\delta - 1)}{\delta + 1 - \gamma(\delta - 1)} U_{n1}^2 - \delta V_{A1}^2 \cos^2 \theta_1 \right] = 0. \end{aligned} \quad (7.3.41)$$

This equation is called the *shock adiabatic* [Anderson, 1963]. It gives the propagation speed of the shock as a function of the shock strength and the upstream parameters. The shock adiabatic is closely analogous to the dispersion relation obtained from linear wave analysis, except the propagation speed now depends on the wave amplitude. As can be seen

from the above equation, the shock adiabatic is a cubic equation in U_{n1}^2 . Thus, three pairs of roots exist, corresponding to three pairs of oppositely propagating shocks. In order of increasing speeds, these shocks are called slow, intermediate, and fast shocks. Since the general solution is a complicated function of the various upstream parameters, it is useful to consider certain limiting cases.

7.3.3 The weak shock limit

To gain insight into the general form of the solutions of the shock adiabatic, we start by considering the weak shock limit, $\delta = 1$. This case is important because each root of the shock adiabatic can be traced back to a root at $\delta = 1$. Setting $\delta = 1$ in equation (7.3.41) gives

$$(U_{n1}^2 - V_{A1}^2 \cos^2 \theta_1)^2 [U_{n1}^2 - V_{S1}^2] - \sin^2 \theta_1 U_{n1}^2 V_{A1}^2 [U_{n1}^2 - V_{A1}^2 \cos^2 \theta_1] = 0 \quad (7.3.42)$$

which simplifies to

$$(U_{n1}^2 - V_{A1}^2 \cos^2 \theta_1) [U_{n1}^4 - U_{n1}^2 (V_{A1}^2 + V_{S1}^2) + V_{S1}^2 V_{A1}^2 \cos^2 \theta_1] = 0 \quad (7.3.43)$$

The above equation is seen to be identical to the dispersion relation, equation (??), for small amplitude waves in an MHD fluid. The shock speed, U_{n1} , plays the same role as the phase speed, v_p , of an MHD wave. This correspondence is not surprising, since for a small amplitude discontinuity the field quantities can be decomposed into an infinite series of plane waves all propagating at the same speed. Explicit solutions for the propagation speeds of the slow (s), intermediate (i), and fast (f) shocks are given below for $\delta = 1$,

$$U_{n1}^2(s) = \frac{1}{2} (V_{A1}^2 + V_{S1}^2) - \frac{1}{2} \left[(V_{A1}^2 + V_{S1}^2)^2 + 4V_{A1}^2 V_{S1}^2 \sin^2 \theta_1 \right]^{1/2}, \quad (7.3.44)$$

$$U_{n1}^2(i) = V_{A1}^2 \cos^2 \theta_1, \quad (7.3.45)$$

$$U_{n1}^2(f) = \frac{1}{2} (V_{A1}^2 + V_{S1}^2) + \frac{1}{2} \left[(V_{A1}^2 - V_{S1}^2)^2 + 4V_{A1}^2 V_{S1}^2 \sin^2 \theta_1 \right]^{1/2}. \quad (7.3.46)$$

Polar plots of the propagation speeds of these shocks are shown as a function of the angle θ between the magnetic field and the shock normal in Figure 7.8. These plots show the shape of the shock fronts that would be produced by a small impulsive disturbance at the origin. As with MHD waves, two cases must be considered, $V_{A1} > V_{S1}$ and $V_{A1} < V_{S1}$. The slow and fast shocks correspond to the slow and fast MHD waves, and the intermediate shock corresponds to the transverse Alfvén wave. The field geometries for the three basic types of small amplitude shocks can be determined by referring to the eigenvector diagrams shown in Figures ?? and ???. In the small amplitude $\delta = 1$ limit, the intermediate shock degenerates to a rotational discontinuity.

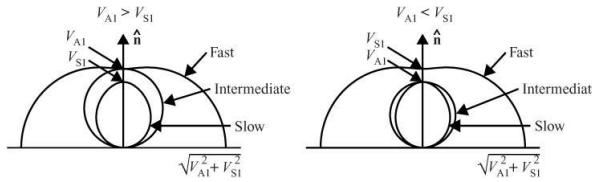


Fig. 7.8 Polar plots of the propagation speeds of the slow, intermediate, and fast shocks in the weak shock limit for the two cases, $V_{A1} > V_{S1}$ and $V_{A1} < V_{S1}$.

7.3.4 Parallel and perpendicular shocks

From simple inspection of the shock adiabatic (7.3.41), it is evident that the equation is greatly simplified whenever either $\theta_1 = 0$ or $\theta_1 = \pi/2$. The case $\theta_1 = 0$ is called a parallel shock, and the case $\theta_1 = \pi/2$ is called a perpendicular shock. For a parallel shock ($\theta_1 = 0$), it is easy to show that the shock adiabatic has two solutions,

$$U_{n1}^2 = \frac{2\delta V_{S1}^2}{\delta + 1 - \gamma(\delta - 1)}, \quad (7.3.47)$$

and

$$U_{n1}^2 = \delta V_{A1}^2. \quad (7.3.48)$$

In the limit $\delta = 1$, the first root corresponds to the slow shock if $V_{A1} > V_{S1}$, and to the fast shock if $V_{A1} < V_{S1}$, and the second root corresponds to the intermediate shock discussed in the previous section. It is interesting to note that the first root has no dependence on the Alfvén speed, and therefore no dependence on the magnetic field strength. The reason that the magnetic field plays no role in this shock can be seen from equation (7.3.23), which relates the tangential components of the magnetic field on the two sides of the shock. For a parallel shock B_{t1} is zero. Provided that $M_{A1}^2 \neq \delta$, which is always the case for the first root, it follows that B_{t2} is zero. Thus, the magnetic field is parallel to the normal on both sides of the shock. Since the normal component of the magnetic field is continuous across the shock ($B_{n1} = B_{n2}$), it follows that the magnetic field is completely unaffected by the shock. The reason the magnetic field plays no role is that for a parallel shock the fluid motion is purely compressional with no velocity component perpendicular to the magnetic field, hence no $\mathbf{U} \times \mathbf{B}$ force. This shock is indistinguishable from a shock in an ordinary gas, and forms because of the non-linear wave steepening effect illustrated in Figure 7.1. Since γ is always greater than one, it is obvious from inspection of equation (7.3.47) that the speed of this shock increases as δ increases, and goes to infinity as the denominator goes to zero at $\delta = \delta_m$, where

$$\delta_m = \frac{\gamma + 1}{\gamma - 1}. \quad (7.3.49)$$

No solution exists for $\delta > \delta_m$. The quantity δ_m is the maximum attainable shock strength,

and therefore represents an upper limit to the density compression ratio, $\delta = \rho_{m2}/\rho_{m1}$. For $\gamma = 5/3$ it is easy to show that $\delta_m = 4$. To show the asymptotic dependence more explicitly as $\delta \rightarrow \delta_m$, it is useful to rewrite the $\delta + 1 - \gamma(\delta - 1)$ term in the denominator of equation (7.3.47) as $(\gamma - 1)(\delta_m - \delta)$. The shock propagation speed is then given by

$$U_{n1}^2 = \frac{2\delta V_{S1}^2}{(\gamma - 1)(\delta_m - \delta)}, \quad (7.3.50)$$

which shows that as $\delta \rightarrow \delta_m$ the speed varies as $1/(\delta_m - \delta)^{1/2}$.

In contrast to the first root, the magnetic field plays an essential role in the second root, $U_{n1}^2 = \delta V_{A1}^2$, which corresponds to the intermediate shock. The reason that the magnetic field plays a role in the propagation of this shock, even though B_{t1} is zero, is that for this root the denominator on the right-hand side of equation (7.3.23) is zero ($M_{A1}^2 = \delta$). This leads to an indeterminacy in B_{t2} , even though B_{t1} is zero. For this shock the magnetic field can be parallel to the normal on the upstream side of the shock, but still have a tangential component on the downstream side. Because of the abrupt appearance of a tangential component to the magnetic field, this shock is sometimes called a *switch-on shock*.

For a perpendicular shock ($\theta = \pi/2$), it is easy to show that the shock adiabatic (equation (7.3.41)) has only one solution

$$U_{n1}^2 = \frac{2\delta[V_{S1}^2 + (\delta - \gamma(\delta - 1)/2)V_{A1}^2]}{(\gamma - 1)(\delta_m - \delta)}, \quad (7.3.51)$$

In the weak shock limit ($\delta = 1$), this solution corresponds to the fast magnetosonic mode. Note that the propagation speed of this shock depends on both the sound speed and the Alfvén speed.

The reason for this mixed dependence is that the plasma and the magnetic field are both compressed as the plasma flows across the shock. Since there is no field line stretching for a perpendicular shock ($\ell_1 = \ell_2$), the frozen field theorem shows that the magnetic field ratio varies in direct proportion to the density compression ratio, $B_2/B_1 = \rho_{m2}/\rho_{m1}$. This same conclusion can be obtained from equation (7.3.23) (note that for $\theta_1 = \pi/2$, $M_{A1} = U_{n1}/V_{A1} \cos \theta_1 \rightarrow \infty$). Since both the plasma pressure and the magnetic field pressure increase across the shock, it is easy to see why the shock propagation speed depends on both the sound speed and the Alfvén speed. Note also that the Alfvén speed increases as the magnetic field strength increases. This amplitude dependence causes the magnetic field to steepen in a way that is analogous to the steepening of a pressure pulse in an ordinary gas.

From inspection of equation (7.3.51), one can see that the propagation speed of a perpendicular shock goes to infinity as the denominator goes to zero, just as is the case in a parallel shock. A perpendicular shock has exactly the same upper limit to the shock strength, δ_m , and asymptotic dependence, $1/(\delta_m - \delta)^{1/2}$, as a parallel shock. Shortly, we will show that this upper limit applies for any orientation of the upstream magnetic field.

7.3.5 The strong shock limit

Since both parallel and perpendicular shocks have an upper limit to the shock strength, it is useful to explore the general solution to the shock adiabatic as δ approaches δ_m . For this purpose, we rewrite the $\delta + 1 - \gamma(\delta - 1)$ term as $(\gamma - 1) \times (\delta_m - \delta)$, and then evaluate the limiting form as $\delta \rightarrow \delta_m$. Expanding the shock adiabatic (7.3.41) in powers of U_{n1} gives the following equation

$$\begin{aligned} U_{n1}^6 - U_{n1}^4 & \left[\frac{2\delta V_{S1}^2 + 2\delta \sin^2 \theta_1 V_{A1}^2 (\delta - \gamma(\delta - 1)/2)}{(\gamma - 1)(\delta_m - \delta)} + 2\delta V_{A1}^2 \cos^2 \theta_1 \right] \\ & + U_{n1}^2 \delta^2 V_{A1}^2 \cos^2 \theta_1 \left[V_{A1}^2 + \frac{4V_{S1}^2}{(\gamma - 1)(\delta_m - \delta)} \right] - \frac{2\delta^3 V_{A1}^4 V_{S1}^2 \cos^4 \theta_1}{(\gamma - 1)(\delta_m - \delta)} = 0. \end{aligned} \quad (7.3.52)$$

As $\delta \rightarrow \delta_m$, it is clear that one of the roots for U_{n1}^2 must approach infinity as $1/(\delta_m - \delta)$. This root is controlled by the U_{n1}^6 and U_{n1}^4 terms in the above equation and is given by

$$U_{n1}^2 = \frac{2\delta [V_{S1}^2 + \sin^2 \theta_1 V_{A1}^2 (\delta - \gamma(\delta - 1)/2)]}{(\gamma - 1)(\delta_m - \delta)} \quad (7.3.53)$$

It is interesting to note that even though the above equation is rigorously valid only for a strong shock, it still reduces exactly to equation (7.3.50) for a parallel shock and exactly to equation (7.3.51) for a perpendicular shock. Note that the shock speed goes to infinity as $1/(\delta_m - \delta)^{1/2}$ for all orientations of the upstream magnetic field, $0 \leq \theta \leq \pi/2$.

The remaining two roots for U_{n1}^2 can be investigated by multiplying equation (7.3.52) by $(\delta_m - \delta)$ and taking the limit as δ approaches δ_m , which gives the equation

$$\begin{aligned} U_{n1}^4 [2V_{S1}^2 + \sin^2 \theta_1 V_{A1}^2 (\delta_m - 1)] - U_{n1}^2 4\delta_m \cos^2 \theta_1 V_{A1}^2 V_{S1}^2 \\ + 2\delta_m^2 V_{A1}^4 V_{S1}^2 \cos^4 \theta_1 = 0. \end{aligned} \quad (7.3.54)$$

The above equation is a quadratic in U_{n1}^2 and in the limit $\delta = \delta_m$ has a discriminant

$$-8\delta_m^2 (\delta_m - 1) \cos^4 \theta_1 \sin^2 \theta_1 V_{A1}^6 V_{S1}^2. \quad (7.3.55)$$

Since δ_m is always greater than one, the discriminant is always negative. Thus, the remaining two roots, which correspond to the slow and intermediate shocks, merge and become complex (i.e., non-propagating) at some intermediate shock strength, δ_c , that is less than δ_m . It follows then that δ_m represents an absolute upper limit to the shock strength for all three types of shocks.

7.3.6 Arbitrary shock strengths and magnetic field orientations

From the above discussion, we can arrive at the following qualitative understanding of the solutions of the shock adiabatic for arbitrary shock strengths and magnetic field orientations.

- (1) The shock adiabatic has three roots, called the slow, intermediate and fast shocks.
- (2) In the weak shock limit, $\delta \approx 1$, the propagation speeds of the three shocks correspond to the propagation speeds of the slow magnetosonic wave, $U_n(s)$, the transverse Alfvén wave, $U_n(i)$, and the fast magnetosonic wave, $U_n(f)$.
- (3) As the shock strength increases, the speeds of the slow and intermediate shocks merge at an intermediate value of the shock strength, δ_c . For any larger value of δ , these shocks no longer exist. Thus, there is an upper limit, δ_m , to the strengths of the slow and intermediate shocks.
- (4) For shock strengths greater than δ_c only one solution exists (the fast shock) and the speed of this shock asymptotically approaches infinity as $1/(\delta_m - \delta)^{1/2}$, where δ_m is given by equation (7.3.49). No solutions exist for shock strengths greater than δ_m .

From these results we can conclude that a plot of the shock speed, U_n , as a function of the shock strength, δ , must have the form shown in Figure 7.9. The exact shape of the plot depends on the specific choices for the upstream parameters, V_{A1} , V_{S1} , and θ_1 , and can only be determined by carrying out a full numerical solution of equation (7.3.41). The plot in Figure 7.9 was made using $V_{A1} = 4$, $V_{S1} = 1$, and $\theta_1 = 60^\circ$.

An important distinction can be made between fast and slow shocks. From equation (7.3.23) it can be shown that $B_{t2} > B_{t1}$ for fast shocks and $B_{t2} < B_{t1}$ for slow shocks. Thus, the magnetic field is rotated away from the normal for a fast shock, $\theta_2 > \theta_1$, and toward the normal, $\theta_2 < \theta_1$, for a slow shock. The corresponding geometries are illustrated in Figure 7.10. The change in the direction of the flow velocity at the shock is caused by the $\mathbf{J} \times \mathbf{B}$ force acting on the fluid at the discontinuity. The current in the discontinuity surface flows upward (out of the paper) in Figure 7.10 for a fast shock, and downward (into the paper) for a slow shock.

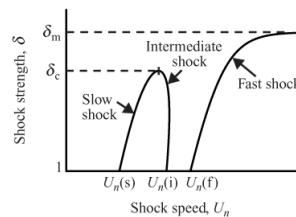


Fig. 7.9

The shock propagation velocity, U_n , as a function of the shock strength, δ , for a constant angle, θ , between the upstream magnetic field direction and the shock normal.

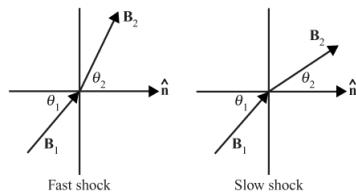


Fig. 7.10 Fast and slow shocks can be distinguished by the deflection of the downstream magnetic field, which is away from the normal for a fast shock, and toward the normal for a slow shock.

7.3.7 Entropy and reversibility

Inspection of equation (7.3.41) clearly shows that for each of the three roots of the shock adiabatic there are two solutions, U_{n1} and $-U_{n1}$. This reversibility arises because the Rankine–Hugoniot equations (7.3.8)–(7.3.14) are time reversible. Since our convention is always to label the upstream side 1 and the downstream side 2, it follows that, upon reversing the direction of propagation, the shock strength changes from δ to $1/\delta$. It can be shown that the shock adiabatic only has solutions for δ values in the range $(1/\delta_m) < \delta < \delta_m$. This includes values less than one. A value of δ less than one implies that the density downstream of the shock is less than the density upstream of the shock. However, such non-compressive shocks can never occur in nature. The reason is that such shocks violate the second law of thermodynamics. To understand the thermodynamic limitations placed on solutions of the Rankine–Hugoniot equations, we must return to a discussion of the compression that occurs as the plasma flows across the shock. Since the kinetic energy decreases as the fluid crosses the shock, it is clear that some of the flow energy is converted to some other form during the compression process, either internal energy (i.e., heat) or magnetic energy. Although the compression in the shock transition region is abrupt, so that there is little time for heat to flow, the adiabatic equation of state (??) cannot be used. The reason is that the fluid is not in a *succession of equilibrium states*, which is a basic assumption in the derivation of the adiabatic equation of state. Such abrupt non-equilibrium processes are always accompanied by dissipation of energy and are irreversible [Sears, 1953]. In a shock the energy dissipation occurs during the sudden compression, which irreversibly converts some of the incoming flow energy into heat. Thus, instead of being determined by the adiabatic equation of state, the downstream parameters are determined by the Rankine–Hugoniot equations. That the adiabatic equation of state is not used may seem surprising, since the energy equation (??) was derived using the adiabatic equation of state. However, the use of the adiabatic equation of state in the derivation of the energy equation was just a convenient way of deriving the internal energy (??) and enthalpy (??). These quantities are state variables and are independent of the process used to go from the initial state to the final state.

To show that the compression that takes place at a shock is an irreversible process, we next compute the change in the entropy as the fluid flows across the shock. The second law of thermodynamics states that the entropy must either increase or be constant and, if

it increases, the process is irreversible. Following the usual procedure used in thermodynamics, the change in the entropy, dS , is computed by integrating $dS = dQ/T$ for some reversible process between the initial and final states. For an ideal gas it can be shown that the change in the specific entropy, $\Delta s = s_2 - s_1$, for an abrupt adiabatic (i.e., no heat loss) compression from some initial pressure and density (P_1, ρ_1) to a final pressure and density (P_2, ρ_2) is given by

$$\Delta s = c_V \ln \left[\frac{P_2}{P_1} \left(\frac{\rho_1}{\rho_2} \right)^\gamma \right], \quad (7.3.56)$$

where c_V is the specific heat capacity per unit volume, and γ is the ratio of the heat capacity at constant pressure to the heat capacity at constant volume [Fetter and Walecka, 1980]. To evaluate the entropy change using the above equation, the first step is to compute the pressure ratio, P_2/P_1 , across the shock as a function of the shock strength, $\delta = \rho_2/\rho_1$. The pressure ratio is most easily obtained from equation (7.3.36), which after substituting for M_A and M_S using equations (7.3.16) and (7.3.17) gives

$$\frac{P_2}{P_1} = 1 + \gamma \frac{U_{n1}^2}{V_{S1}^2} \frac{(\delta - 1)}{\delta} \left[1 - \frac{\delta V_{A1}^2 [(\delta + 1)U_{n1}^2 - 2\delta V_{A1}^2 \cos^2 \theta_1]}{2(U_{n1}^2 - \delta V_{A1}^2 \cos^2 \theta_1)^2} \sin^2 \theta_1 \right]. \quad (7.3.57)$$

To compute the pressure ratio we must solve the shock adiabatic (7.3.41) for the shock speed U_{n1} and then substitute the shock speed into the above equation. Once the pressure ratio is known, the entropy change can be computed from equation (7.3.56), noting that $\rho_2/\rho_1 = \delta$. Usually this computation must be done numerically. However, in certain cases it is possible to carry out the calculation analytically. For example, in the strong shock limit, where the shock speed varies as $U_{n1} \propto 1/(\delta_m - \delta)^{1/2}$ the second term in the brackets becomes negligibly small as $\delta \rightarrow \delta_m$. The pressure ratio is then determined by the term in front of the brackets, which goes to plus infinity as $1/(\delta_m - \delta)$. Thus, for a strong compressible shock the entropy change is clearly positive, which shows that the shock can exist, and that the process is irreversible. In the weak shock limit, it is evident from equation (7.3.57) that the pressure ratio goes to one as $\delta \rightarrow 1$. The entropy change then goes to zero. Thus, in the weak shock limit the shock propagation is reversible. This is to be expected, since as was shown earlier, small amplitude MHD waves propagate with essentially no energy dissipation. That the entropy is positive for all values of δ greater than one, and negative for all values of δ less than one, is difficult to prove in general but is easy to demonstrate for specific cases. For example, for a parallel shock, $\theta_1 = 0$, the shock speed given by equation (7.3.50) can be substituted directly into equation (7.3.57) to give the following equation for the pressure ratio as a function of the shock strength

$$\frac{P_2}{P_1} = \frac{\delta \delta_m - 1}{\delta_m - \delta} \quad (7.3.58)$$

where we have eliminated γ using $\gamma = (\delta_m + 1)/(\delta_m - 1)$, which is derivable from equation (7.3.49). A plot of the corresponding entropy change, Δs , computed from equation (7.3.56) for $\theta_1 = 0$, is shown in Figure 7.11 as a function of the shock strength, δ . Also shown is the

entropy change computed from equations (7.3.51), (7.3.56) and (7.3.57) for $\theta_1 = \pi/2$. As can be seen in both cases the entropy change goes to zero at $\delta = 1$, and is positive for $\delta > 1$ and negative for $\delta < 1$. This means that these shocks can only occur if the shock strength, $\delta = \rho_2/\rho_1$, is greater than one (i.e., only if the fluid is compressed).

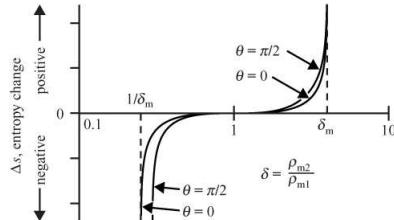


Fig. 7.11 A plot of the entropy change, Δs , for a fast shock as a function of the shock strength, $\delta = \rho_2/\rho_1$, for a specific choice of upstream parameters ($V_{A1} = 40 \text{ km s}^{-1}$, $V_{S1} = 30 \text{ km s}^{-1}$, $\theta = 0$, and $\gamma = 5/3$).

7.3.8 Non-linear evolution and stability

The Rankine-Hugoniot equations implicitly assume that the flow through the shock is in steady state, i.e., all of the time dependent terms in the MHD equations are zero. However, one must ask whether such a steady-state solution can actually evolve from some arbitrary initial condition. As discussed in the beginning of this chapter, for a discontinuity to exist, the initial disturbance must undergo the non-linear process known as wave steepening. If an arbitrary initial disturbance does not undergo wave steepening then a shock cannot form, even though a formal solution to the jump conditions may exist. The question of whether an arbitrary initial disturbance evolves into a shock is difficult to answer in general, since it involves the full nonlinear solution of the time dependent MHD equations. However, on a qualitative level, it is easy to see that both the slow and fast magnetosonic modes have the proper non-linear characteristics to lead to wave steepening. That this is the case can be seen from a simple inspection of equations (??), (??), (??) and (??), which shows that the phase velocities of both the slow and fast magnetosonic modes increase as the pressure and/or the magnetic field strength increase (i.e., the larger amplitude part of the disturbance tends to overtake the smaller amplitude part of the disturbance). Thus, slow and fast shocks arise from steepening of the slow and fast magnetosonic modes, a conclusion that can be verified by numerically solving the time dependent MHD equations. On the other hand, the transverse Alfvén mode, which is linked to the intermediate shock, does not tend to steepen, since the phase velocity, which is determined by the zero-order background magnetic field, is independent of the magnetic field perturbation, which is perpendicular to the zero-order magnetic field (see Figure ??). As mentioned earlier, for small amplitudes and $\theta = 0$, the intermediate shock degenerates to a rotational discontinuity. Thus, based on these simple considerations, there are good reasons to believe that a transverse Alfvén wave cannot evolve into an intermediate shock.

An alternative approach to answering the question of whether a shock can exist is to

test the stability of the shock by introducing a small amplitude perturbation in the shock boundary, and then carrying out a time dependent linear analysis to determine whether the resulting perturbation grows in amplitude. If the perturbation grows, the steady-state solution is unstable and cannot occur in nature, since if a shock somehow formed, it would quickly break up into some other type of non-steady flow. This type of stability analysis is difficult, and will not be pursued here. However, the issues involved have been studied in considerable detail by several authors [Anderson, 1963; Kantrowitz and Petschek, 1966; Kennel *et al.*, 1989], and the conclusions are that the slow and fast shocks are stable, but the intermediate shock is not. Thus, of the three possible solutions of the shock adiabatic, only the fast and slow shocks are believed actually to occur in nature.

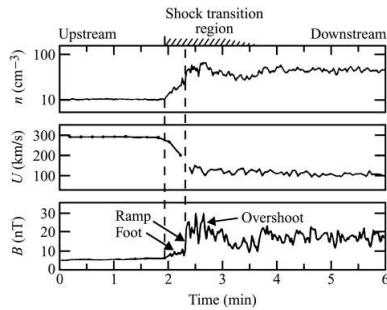


Fig. 7.12 An example of a bow shock crossing observed at a geocentric radial distance of $15.4 R_E$ upstream of the Earth [from Scudder *et al.*, 1986].

7.3.9 Observations of MHD shocks

MHD shocks have been observed in a large variety of astrophysical, geophysical, and laboratory plasmas. Shocks are known to be produced by supernova explosions, by the outflow of hot gases from certain types of stars, by solar flares, and in the solar wind upstream of planetary magnetospheres. Of these, the shocks that form in the solar wind upstream of planetary magnetospheres have received considerable study and are worthy of some brief comments. This type of shock is often called the bow shock, in analogy with the bow wave of a ship. An example of a bow shock observed upstream of the Earth's magnetosphere is shown in Figure 7.12. In this case the angle between the shock normal and the magnetic field, $\theta_1 = 76^\circ$, is quite large, and the upstream flow speed, $U_1 = 294 \text{ km s}^{-1}$, is much greater than either the Alfvén speed or the sound speed, which are $V_{A1} = 37.8 \text{ km s}^{-1}$, and $V_{S1} = 58.6 \text{ km s}^{-1}$. Therefore, this is a strong nearly perpendicular fast shock. Because the solar wind velocity, typically about 400 km s^{-1} , is usually much greater than either the Alfvén speed or the sound speed, planetary bow shocks are almost always strong fast shocks. The abrupt increases in the number density, n , and the magnetic field strength, B , and the decrease in the flow velocity, U , at the shock are clearly evident. Also, note that the ratio of the upstream number density to the downstream number density (i.e., $n_1/n_2 = \rho_{m1}/\rho_{m2}$) is almost exactly the theoretical limit for a strong shock, i.e., $\delta_m = 4$.

As one can see from Figure 7.12, the shock transition region has considerable structure extending over a period of nearly two minutes. This time span corresponds to a thickness of several hundred kilometers [Scudder *et al.*, 1986]. When standing bow shocks were first discovered in the solar wind, one of the main questions was what controlled the thickness of the shock, which is much smaller than the mean-free path. Typical mean-free paths in the solar wind are a substantial fraction of an astronomical unit (i.e., $\sim 10^8$ km). It is now known that various characteristic length scales of the plasma, such as the electron and ion cyclotron radii, play a crucial role in determining the thickness and structure of the shock. Since the cyclotron radius of the ions is much larger than the cyclotron radius of the electrons, when the incoming plasma encounters the main jump in the magnetic field, labeled “ramp” in Figure 7.12, the electrons and ions respond quite differently. Because of their small cyclotron radius relative to the field gradients the electrons are essentially carried along with the fluid flow (i.e., at the $\mathbf{E} \times \mathbf{B}$ drift velocity) while conserving the 1st adiabatic invariant. Because of their much larger cyclotron radius, the ions are strongly deflected by the $\mathbf{v} \times \mathbf{B}$ force and move on complicated non-adiabatic trajectories that are quite different than that of the electrons. In some cases the ions can even be reflected back into the upstream region, where they are accelerated by the solar wind electric field and carried back into the shock again by the $\mathbf{E} \times \mathbf{B}$ drift. Such reflected ions are responsible for the region labeled “foot” that lies ahead of the main ramp in the magnetic field in Figure 7.12. The resulting differential motion between the ions and the electrons leads to a charge separation and currents that are self-consistently responsible for the electric and magnetic fields in the shock transition region. As the plasma flows through this system of quasi-static electric and magnetic fields, the distribution functions are drastically altered, especially for the ions, resulting in a hot highly non-thermal plasma made up of ring-like and shell-like particle distributions. These non-thermal particle distributions are unstable and rapidly lead to the generation of plasma wave turbulence via various plasma instabilities, some of which are discussed in the next two chapters. This turbulence then acts to scatter the particles and drive the distribution functions toward thermal equilibrium, effectively playing the role of collisions. The net result is the conversion of the directed kinetic energy of the incoming plasma into random thermal energy, very similar to what occurs in a collision dominated shock.

To illustrate the types of plasma waves that occur in a collisionless shock, an electric field spectrogram obtained at Jupiter’s bow shock is shown in Figure 7.13. The narrowband emissions at about 6 kHz are electron plasma oscillations excited by a beam of electrons that escapes into the region upstream of the shock. The intense broadband electric field noise starting at about 30 seconds and extending into the downstream region is plasma wave turbulence excited by unstable particle distributions in the shock transition region. It is this turbulence that plays a key role in producing the irreversible dissipation that converts the incoming flow energy into random thermal energy.

The thickness of the shock and the detailed substructure within the shock depend on many factors, the most important being the angle θ_1 between the upstream magnetic field and the shock normal, and the two upstream Mach numbers, M_{A1} and M_{S1} . The transition region of a quasi-perpendicular shock is usually thin and well defined. The transition region of a quasi-parallel shock is usually more complicated, and often appears thick and poorly

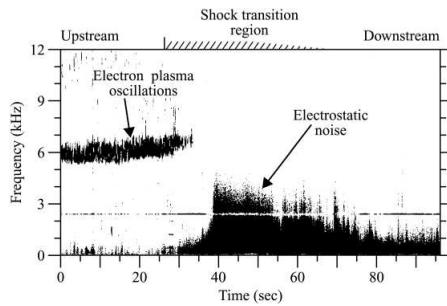


Fig. 7.13 An electric field spectrogram showing the plasma waves observed by the Voyager 1 spacecraft during a crossing of Jupiter's bow shock [Gurnett and Scarf, 1983].

defined in the magnetic field data, but less so in the plasma data [see Scudder *et al.*, 1984]. The physical parameters that control the thickness are mainly the cyclotron radius of the ions and a characteristic length called the electron inertial length, which is the spatial scale over which the electrons can respond to the electrostatic fields in the shock. For a more detailed discussion of the structure of the shock transition layer, see Kivelson and Russell [1995] and Baumjohann and Treumann [1997].

Problems

- 7.1. By eliminating $[(M_{A2}^2/M_{S2}^2) - (M_{A1}^2/M_{S1}^2)]$ from equations (7.3.36) and (7.3.38) and eliminating M_{A2}^2 using equation (7.3.39), show that the Alfvén Mach number, M_{A1} , is given by

$$(M_{A1}^2 - \delta)^2 \left[M_{S1}^2 - \frac{2\delta}{\delta + 1 - \gamma(\delta - 1)} \right] - \delta \tan^2 \theta_1 M_{S1}^2 \left[\frac{2\delta - \gamma(\delta - 1)}{\delta + 1 - \gamma(\delta - 1)} M_{A1}^2 - \delta \right] = 0.$$

- 7.2. By making the substitution, $M_{A1} = U_{n1}/(V_{A1} \cos \theta_1)$, show that the above equation simplifies to

$$(U_{n1}^2 - \delta V_{A1}^2 \cos^2 \theta_1)^2 \left[U_{n1}^2 - \frac{2\delta V_{S1}^2}{\delta + 1 - \gamma(\delta - 1)} \right] - \delta \sin^2 \theta_1 U_{n1}^2 V_{A1}^2 \left[\frac{2\delta - \gamma(\delta - 1)}{\delta + 1 - \gamma(\delta - 1)} U_{n1}^2 - \delta V_{A1}^2 \cos^2 \theta_1 \right] = 0.$$

- 7.3. Show that there is only one solution to the shock adiabatic in the limit $\delta \rightarrow \delta_m$. Hint: After finding the root given by equation (7.3.52), show that in the limit $\delta \rightarrow \delta_m$ the remaining two roots to the shock adiabatic are solutions to the equation

$$U_n^4 \left[2V_{S1}^2 + \sin^2 \theta V_{A1}^2 (\delta_m - 1) \right] - U_{n1}^2 4\delta_m \cos^2 \theta_1 V_{A1}^2 V_{S1}^2 + 2\delta^2 V_{A1}^4 V_{S1}^2 \cos^4 \theta_1 = 0,$$

and that this quadratic equation has a discriminant

$$-8\delta_m^2(\delta_m - 1) \cos^4 \theta_1 \sin^2 \theta_1 V_{A1}^6 V_{S1}^2,$$

which is always negative for $\delta_m > 1$.

- 7.4. By numerically solving the shock adiabatic (7.3.41), show that the three roots for U_{n1} have the qualitative form given in Figure 7.9. Make a plot of U_{n1} versus θ for each of the three roots. For representative parameters use $V_{A1} = 40 \text{ km s}^{-1}$, $V_{S1} = 30 \text{ km s}^{-1}$, and $\theta = 45^\circ$.
- 7.5. By substituting $M_S = U_n \sqrt{\rho_m/\gamma P}$ for M_{S2} and M_{S1} into equation (7.3.36) and substituting $M_{A1}^2 = \mu_0 \rho_{m1} U_{n1}^2 / B_n^2$, show that the downstream to upstream pressure ratio can be written

$$\frac{P_2}{P_1} = 1 + \gamma \frac{U_{n1}^2}{V_{S1}^2} \frac{(\delta - 1)}{\delta} \left[1 - \frac{\delta V_{A1}^2 [(\delta + 1)U_{n1}^2 - 2\delta V_{A1}^2 \cos^2 \theta_1]}{2(U_{n1}^2 - \delta V_{A1}^2 \cos^2 \theta_1)^2} \sin^2 \theta_1 \right].$$

- 7.6. Show that for a parallel ($\theta_1 = 0$) fast mode shock the downstream to upstream pressure ratio simplifies to

$$\frac{P_2}{P_1} = \frac{\delta \delta_m - 1}{\delta_m - \delta},$$

where $\gamma = (\delta_m + 1)/(\delta_m - 1)$ has been substituted for γ . Plot P_2/P_1 as a function of δ for $\delta_m = 4$.

References

- Anderson, J. E. 1963. *Magnetohydrodynamic Shock Waves*. MIT Press: Cambridge, MA, p. 21.
- Baumjohann, W., and Treumann, R. A. *Basic Space Plasma Physics*. World Scientific Publishing: London, 1997, Chapter 8.
- Currie, I. G. 1974. *Fundamental Mechanics of Fluids*. McGraw-Hill: New York, p. 337.
- Fetter, A. L., and Walecka, J. D. 1980. *Theoretical Mechanics of Particles and Continua*. McGraw-Hill: New York, p. 345.
- Gurnett, D. A., and Scarf, F. L. 1983. Plasma waves in the Jovian magnetosphere. In *Physics of the Jovian Magnetosphere*, ed. A. J. Dessler. Cambridge University Press: Cambridge, p. 292.
- de Hoffmann, F., and Teller, E. 1950. Magneto-hydrodynamic shocks. *Phys. Rev.* **80**, 691–703.
- Kantrowitz, A., and Petschek, H. E. 1966. MHD characteristics and shock waves. In *Plasma Physics in Theory and Application*, ed. W. B. Kunkel. McGraw-Hill: New York, pp. 148–206.
- Kennel, C. F., Blandford, R. D., and Coppi, P. 1989. MHD intermediate shock discontinuities. Part 1. Rankine–Hugoniot conditions. *J. Plasma Phys.* **43** (2), 299–319.
- Kivelson, M. G., and Russell, C. T. *Introduction to Space Physics*. Cambridge University Press: Cambridge, 1995, Chapter 5.
- Landau, L. D., and Lifshitz, E. M. 1960. *Electrodynamics of Continuous Media*. Pergamon Press: New York, p. 224.
- Parks, G. K. 1991. *Physics of Space Plasmas: An Introduction*. Addison-Wesley: New York, p. 7.
- Scudder, J. D., Burlaga, L. F., and Greenspan, E. W. 1984. Scale lengths at quasi-parallel shocks. *J. Geophys. Res.* **89**, 7545–7550.
- Scudder, J. D., Mangeney, A., Lacombe, C., Harvey, C. C., Aggson, T. L., Anderson, R. R., Gosling, J. T., Paschmann, G., and Russell, C. T. 1986. The resolved layer of a collisionless, high β , supercritical, quasi-perpendicular shock wave: 1. Rankine–Hugoniot geometry, currents and stationarity. *J. Geophys. Res.* **91**, 11019–11052.
- Sears, F. W. 1953. *Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics*. Addison-Wesley: Reading, MA, p. 3.
- Smith, E. J. 1973. Identification of interplanetary tangential and rotational discontinuities. *J. Geophys. Res.* **78**, 2054–2063.
- Sonnerup, B. U. O., Paschmann, G., Papamastorakis, I., Sckopke, N., Haerendel, G., Bame, S. J., Asbridge, J. R., Gosling, J. T., and Russell, C. T. 1981. Evidence for magnetic field reconnection at the Earth's magnetopause. *J. Geophys. Res.* **86**, 10049–10067.

Further reading

- Anderson, J. E. *Magnetohydrodynamic Shock Waves*. MIT Press: Cambridge, MA, 1963, Chapter 2.
- Boyd, T. J. M., and Sanderson, J. J. *The Physics of Plasmas*. Cambridge University Press: Cambridge, 2003, Chapter 5.
- Landau, L. D., and Lifshitz, E. M. *Electrodynamics of Continuous Media*. Pergamon Press: New York, 1960, Chapter 8.
- Parks, G. K. *Physics of Space Plasmas*. Addison-Wesley: Redwood City, CA, 1991, Chapter 10.
- Tidman, D. A., and Krall, N. A. *Shock Waves in Collisionless Plasmas*. Wiley: New York, 1971, Chapter 1.

Electrostatic waves in a hot unmagnetized plasma

In this chapter we investigate the propagation of small amplitude waves in a hot unmagnetized plasma. Because of the shortcomings of the moment equations, the approach used is to solve the Vlasov equation directly using a linearization procedure similar to that used in the analysis of cold plasmas. Although both electromagnetic and electrostatic solutions exist, the discussion in this chapter is limited to solutions that are purely electrostatic, i.e., the electric field is derivable from the gradient of a potential, $\mathbf{E} = -\nabla\Phi$. Electromagnetic solutions are discussed in the next chapter.

From Faraday's law it is easily verified that electrostatic waves have no magnetic component. This greatly simplifies the Vlasov equation by eliminating the $\mathbf{v} \times \mathbf{B}$ force. For electrostatic waves, it is usually easier to solve for the potential rather than for the electric field. Therefore, in the following analysis, the electric field is replaced by $\mathbf{E} = -\nabla\Phi$, and the potential is calculated from Poisson's equation, $\nabla^2\Phi = -\rho_q/\epsilon_0$.

8.1 The Vlasov approach

In an initial attempt to analyze the problem, we assume that normal modes of the form $\exp(-i\omega t)$ exist and represent them by using Fourier transforms, following the same basic procedure used in Chapter ???. This is the approach used by Vlasov [1945], who first considered this problem. As we will see, the Vlasov approach encounters a mathematical difficulty that can only be resolved by reformulating the problem in terms of Laplace transforms.

As we have done before, we start by only considering the motion of the electrons. The ions are considered to be immobile, with the same zero-order number density, n_0 , as the electrons. The effects of ion motions will be considered later. The equations that must be solved then consist of Vlasov's equation (with $q = -e$, $\mathbf{E} = -\nabla\Phi$ and $\mathbf{B} = 0$)

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f + \frac{e}{m} \nabla\Phi \cdot \nabla_{\mathbf{v}} f = 0, \quad (8.1.1)$$

and Poisson's equation

$$\nabla^2\Phi = -\frac{\rho_q}{\epsilon_0} = -\frac{e}{\epsilon_0} \left[n_0 - \int_{-\infty}^{\infty} f d^3v \right]. \quad (8.1.2)$$

Following the usual linearization procedure, we assume that the electron velocity distri-

bution function, $f(\mathbf{v})$, consists of a constant uniform zero-order distribution, $f_0(\mathbf{v})$, plus a small first-order perturbation, $f_1(\mathbf{v})$

$$f(\mathbf{v}) = f_0(\mathbf{v}) + f_1(\mathbf{v}). \quad (8.1.3)$$

Similarly, we write $\Phi = \Phi_1$. Linearizing equations (8.1.1) and (8.1.2) and noting that $\int f_0 d^3 v = n_0$, we obtain for the first-order equations,

$$\frac{\partial f_1}{\partial t} + \mathbf{v} \cdot \nabla f_1 + \frac{e}{m} \nabla \Phi_1 \cdot \nabla_{\mathbf{v}} f_0 = 0, \quad (8.1.4)$$

and

$$\nabla^2 \Phi_1 = \frac{e}{\epsilon_0} \int_{-\infty}^{\infty} f_1(\mathbf{v}) d^3 v. \quad (8.1.5)$$

Next, we attempt to solve these equations using the usual Fourier transform approach. Since no preferred direction exists, the z axis can be aligned parallel to \mathbf{k} . After dropping the subscript 1 on the first-order terms and making the usual operator substitutions ($\partial/\partial t \rightarrow -i\omega$ and $\nabla \rightarrow i\mathbf{k}$), it is easy to see that the Fourier transformed equations are

$$-i\omega \tilde{f} + ikv_z \tilde{f} + i \frac{e}{m} k \tilde{\Phi} \frac{\partial f_0}{\partial v_z} = 0, \quad (8.1.6)$$

and

$$k^2 \tilde{\Phi} = -\frac{e}{\epsilon_0} \int_{-\infty}^{\infty} \tilde{f}(\mathbf{v}) d^3 v. \quad (8.1.7)$$

Solving for \tilde{f} from equation (8.1.6),

$$\tilde{f} = \frac{-1}{(kv_z - \omega)} \frac{e}{m} k \tilde{\Phi} \frac{\partial f_0}{\partial v_z}, \quad (8.1.8)$$

and substituting into equation (8.1.7) gives

$$k^2 \tilde{\Phi} = \frac{e^2}{\epsilon_0 m} k \tilde{\Phi} \int_{-\infty}^{\infty} \frac{(\partial f_0 / \partial v_z)}{(v_z - \omega/k)} d^3 v. \quad (8.1.9)$$

Rearranging the terms and factoring out $\tilde{\Phi}$ then gives the homogeneous equation

$$\left[1 - \frac{e^2}{\epsilon_0 m k^2} \int_{-\infty}^{\infty} \frac{\partial f_0 / \partial v_z}{(v_z - \omega/k)} d^3 v \right] \tilde{\Phi} = 0. \quad (8.1.10)$$

For the potential $\tilde{\Phi}$ to have a non-trivial solution, the term in the brackets must be zero, which gives the dispersion relation

$$D(k, \omega) = 1 - \frac{e^2}{\epsilon_0 m k^2} \int_{-\infty}^{\infty} \frac{\partial f_0 / \partial v_z}{(v_z - \omega/k)} d^3 v = 0. \quad (8.1.11)$$

At this point, it is useful to define a new normalized one-dimensional distribution function

$$F_0(v_z) = \frac{1}{n_0} \int_{-\infty}^{\infty} f_0(\mathbf{v}) dv_x dv_y. \quad (8.1.12)$$

The function $F_0(v_z)$ is sometimes called the reduced distribution function. Note that by dividing by n_0 the reduced distribution function is normalized such that $\int F_0(v_z) dv_z = 1$. Recognizing that $\omega_p^2 = n_0 e^2 / (\epsilon_0 m)$, the dispersion relation (8.1.11) can be written

$$D(k, \omega) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{(v_z - \omega/k)} dv_z = 0. \quad (8.1.13)$$

In some applications it is useful to put the dispersion relation into a slightly different form by integrating by parts once, which gives

$$\int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{(v_z - \omega/k)} dv_z = \left[\frac{F_0}{(v_z - \omega/k)} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \frac{F_0}{(v_z - \omega/k)^2} dv_z. \quad (8.1.14)$$

Since F_0 goes to zero at $v_z = \pm\infty$, the first term in the integration by parts vanishes, so the dispersion relation becomes

$$D(k, \omega) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{F_0}{(v_z - \omega/k)^2} dv_z = 0. \quad (8.1.15)$$

Both of the above forms of the dispersion relation suffer from a serious problem. Because the denominator goes to zero at $v_z = \omega/k$, the integrals do not converge unless F_0 and $\partial F_0 / \partial v_z$ are zero at $v_z = \omega/k$. Physically, the dispersion relation exists only if there are no particles moving with a velocity equal to the phase velocity of the wave. For the moment, we restrict our consideration to distribution functions for which F_0 and $\partial F_0 / \partial v_z$ are zero at $v_z = \omega/k$. In the next section we will return to the issue of how to deal with a distribution function that is not zero at $v_z = \omega/k$.

Since the charge only enters into the dispersion equation as a squared, e^2 , term in ω_p^2 , it is easy to see that ion motions can be incorporated by replacing ω_p and F_0 by ω_{ps} and F_{s0} , respectively, and summing over all species. For an arbitrary number of species, the general form of the dispersion relation is then

$$D(k, \omega) = 1 - \sum_s \frac{\omega_{ps}^2}{k^2} \int_{-\infty}^{\infty} \frac{\partial F_{s0} / \partial v_z}{(v_z - \omega/k)} dv_z = 0, \quad (8.1.16)$$

or equivalently,

$$D(k, \omega) = 1 - \sum_s \frac{\omega_{ps}^2}{k^2} \int_{-\infty}^{\infty} \frac{F_{s0}}{(v_z - \omega/k)^2} dv_z = 0. \quad (8.1.17)$$

These forms will be useful later when we consider ion effects.

8.1.1 The dispersion relation

To get back to familiar ground (Langmuir oscillations), we next evaluate the dispersion relation assuming that the phase velocity is much greater than the electron thermal velocity, i.e., $(\omega/k)^2 \gg \langle v_z^2 \rangle$. Then the integrand in equation (8.1.15) can be expanded in powers of the small quantity $k v_z / \omega$

$$\frac{1}{(1 - k v_z / \omega)^2} = 1 + 2 \left(\frac{k v_z}{\omega} \right) + 3 \left(\frac{k v_z}{\omega} \right)^2 + \dots \quad (8.1.18)$$

This expansion is sometimes called the high-phase velocity expansion since $\omega/k \gg v_z$. Using this expansion, the dispersion relation can be written

$$D(k, \omega) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \left[1 + 2 \left(\frac{k v_z}{\omega} \right) + 3 \left(\frac{k v_z}{\omega} \right)^2 + \dots \right] F_{e0} \, dv_z. \quad (8.1.19)$$

If we assume that there is no zero-order current, the first moment is zero, $\int v_z F_{e0} \, dv_z = 0$. Retaining terms up to the second moment, $\langle v_z^2 \rangle = \int v_z^2 F_{e0} \, dv_z$, then gives

$$D(k, \omega) = 1 - \frac{\omega_p^2}{\omega^2} \left(1 + 3 \frac{k^2}{\omega^2} \langle v_z^2 \rangle \right) = 0, \quad (8.1.20)$$

or

$$\omega^2 = \omega_p^2 \left(1 + 3 \frac{k^2 \langle v_z^2 \rangle}{\omega^2} \right). \quad (8.1.21)$$

This equation is most conveniently solved by using the method of successive approximations. Since $(\omega/k)^2 \gg \langle v_z^2 \rangle$, the first approximation is

$$\omega^2 = \omega_p^2. \quad (8.1.22)$$

Substituting $\omega^2 = \omega_p^2$ into the right-hand side of equation (8.1.21) then gives the second approximation

$$\omega^2 = \omega_p^2 + 3 \langle v_z^2 \rangle k^2. \quad (8.1.23)$$

If the velocity distribution function is a Maxwellian, it is easy to show that $\langle v_z^2 \rangle = \kappa T_e / m_e$. The dispersion relation then becomes

$$\omega^2 = \omega_p^2 + 3 \left(\frac{\kappa T_e}{m_e} \right) k^2. \quad (8.1.24)$$

Note that if the temperature is zero, we recover the cold plasma result, $\omega^2 = \omega_p^2$, which is a pure oscillation at the electron plasma frequency.

Since $C_e^2 = \kappa T_e / m_e$, the dispersion relation given by equation (8.1.24) is seen to be

identical to the Bohm–Gross dispersion relation derived in Chapter ?? using the adiabatic equation of state

$$\omega^2 = \omega_p^2 + \gamma_e C_e^2 k^2, \quad (8.1.25)$$

provided we use $\gamma_e = 3$. The fact that $\gamma_e = 3$ implies that the electrons move with one degree of freedom (i.e., $f = 1$ in equation (??)), a result that is not obvious from the moment equation approach.

8.1.2 The dispersion relation

Next, we consider an electron velocity distribution function consisting of two or more beams of cold electrons moving parallel to the z axis. Letting n_j and V_j be the number density and velocity of the j th beam, the zero-order distribution function can be represented by a product of delta functions,

$$f_{e0}(\mathbf{v}) = \sum_j n_j \delta(v_x) \delta(v_y) \delta(v_z - V_j). \quad (8.1.26)$$

Using equation (8.1.12), the reduced distribution function is given by

$$F_{e0}(v_z) = \frac{1}{n_0} \int_{-\infty}^{\infty} f_{e0}(\mathbf{v}) dv_x dv_y = \sum_j \frac{n_j}{n_0} \delta(v_z - V_j), \quad (8.1.27)$$

where $n_0 = \sum_j n_j$. Since $F_{e0}(v_z)$ is zero except when $v_z = V_j$, equation (8.1.15) can be used and yields

$$D(k, \omega) = 1 - \frac{\omega_p^2}{k^2} \sum_j \frac{n_j}{n_0} \int_{-\infty}^{\infty} \frac{\delta(v_z - V_j)}{(v_z - \omega/k)^2} dv_z = 0. \quad (8.1.28)$$

Evaluating the integral and defining $\omega_{pj}^2 = n_j e^2 / (\epsilon_0 m_e)$ as the plasma frequency of the j th beam gives the dispersion relation

$$D(k, \omega) = 1 - \sum_j \frac{\omega_{pj}^2}{k^2} \frac{1}{(V_j - \omega/k)^2} = 0. \quad (8.1.29)$$

Note that since the charge enters into the dispersion relation as e^2 , it does not matter if the beam particles are positively or negatively charged. Therefore, equation (8.1.29) can also be used to analyze electron-ion plasmas. The only requirement is that the total zero-order charge density be zero (i.e., electrical neutrality).

The roots of the dispersion relation (8.1.29) are most conveniently analyzed by writing equation (8.1.29) as

$$H(\omega/k) = \sum_j \frac{\omega_{pj}^2}{(V_j - \omega/k)^2} = k^2, \quad (8.1.30)$$

which for two beams has the form shown in Figure 8.1. As one can see, each beam introduces two roots. If $k > k_c$, the roots are all real. These roots correspond to electrostatic oscillations that move along the beam at phase velocities near the beam velocities. If $k < k_c$, the roots that were located between V_1 and V_2 become imaginary.

Since H is real, these roots always occur as complex conjugate pairs, $\omega_- = \omega_r - i\gamma$ and $\omega_+ = \omega_r + i\gamma$, where γ is assumed to be positive. The root ω_+ corresponds

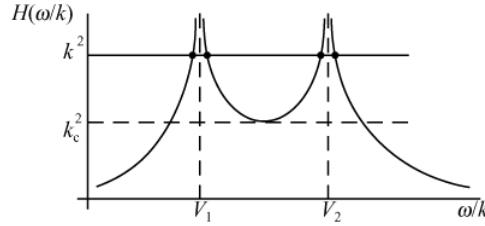


Fig. 8.1 A plot of $H(\omega/k)$ as a function of the phase velocity ω/k for two equal density beams of velocities V_1 and V_2 .

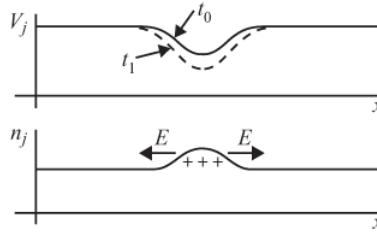


Fig. 8.2 An illustration showing that a perturbation at time t_0 causes a charge imbalance to develop that leads to an even larger perturbation (dashed line) at a later time t_1 .

to a wave that is growing exponentially in time

$$E = E_0 e^{-i\omega_r t} e^{\gamma t}. \quad (8.1.31)$$

The quantity γ is called the growth rate. The above analysis shows that for sufficiently long wavelengths, $k < k_c$, a cold multi-beam plasma is always unstable. This instability is called the *two-stream instability*, since the interaction involves two or more streams of particles.

The physical basis for the two-stream instability is easy to analyze. For a steady-state beam streaming through a fixed background of neutralizing particles, conservation of particles implies that the product of the number density of the beam and the velocity of the beam must be constant:

$$n_j V_j = \text{constant}. \quad (8.1.32)$$

If some perturbation causes a reduction in the velocity at time t_0 , the number density must increase as illustrated in Figure 8.2. Assuming that the beam consists of positive charges (the same basic argument holds for negative charges), the resulting increase in the number

density of the beam causes an excess of positive charge. This positive charge produces an electric field that is directed away from the region of positive charge, which in turn causes a further decrease in the beam velocity, thereby causing the initial perturbation to grow in amplitude, as at time t_1 . It is this feedback, via the electric field, that causes the two-stream instability.

8.1.2.1 Two oppositely directed beams of equal density

For the special case of two beams of equal density ($n_1 = n_2 = n_0$) and equal but opposite velocities ($V_1 = -V_2 = V$), it is possible to obtain a simple analytical expression for the growth rate. The dispersion relation in this case is

$$D(k, \omega) = 1 - \frac{\omega_p^2}{(\omega - kV)^2} - \frac{\omega_p^2}{(\omega + kV)^2} = 0, \quad (8.1.33)$$

where $\omega_p^2 = n_0 e^2 / (\epsilon_0 m)$. The above equation, which is quadratic in ω^2 , has four roots given by

$$\omega^2 = \omega_p^2 + k^2 V^2 \pm (\omega_p^4 + 4k^2 V^2 \omega_p^2)^{1/2}. \quad (8.1.34)$$

It is easy to show that if $k > \sqrt{2}\omega_p/V = k_c$, then ω is real and all of the modes are purely oscillatory. However, if $k < k_c$, two of the oscillatory modes are replaced by an exponentially growing mode and an exponentially decaying mode. Writing $\omega = \pm i\gamma$ for these two roots, the growth rate of the unstable mode is given by

$$\gamma = [(\omega_p^4 + 4k^2 V^2 \omega_p^2)^{1/2} - (\omega_p^2 + k^2 V^2)]^{1/2}. \quad (8.1.35)$$

The maximum growth rate is easily shown to be $\gamma_{\max} = \omega_p/2$ and occurs at $k_{\max} = \sqrt{3/8}k_c$. A plot of the growth rate as a function of wave number is shown in Figure 8.3. Note that in the unstable region, the fields grow exponentially in time with no oscillations ($\omega = 0$). The absence of an oscillatory component is a consequence of the choice of the reference frame which, in this case, is such that the two electron beams have equal and opposite velocities (i.e., $V_1 = -V_2$). In a frame of reference moving at a velocity V' , the fields would oscillate at a frequency $\omega' = kV'$, where kV' is the Doppler shift.

8.1.2.2 The weak-beam approximation

In various applications, one sometimes encounters situations where a very low-density electron beam of plasma frequency ω_b streams through a background of electrons initially at rest with a plasma frequency ω_p . The dispersion relation in this case is easily shown to be

$$D(k, \omega) = 1 - \frac{\omega_p^2}{\omega^2} - \frac{\omega_b^2}{(kV - \omega)^2} = 0, \quad (8.1.36)$$

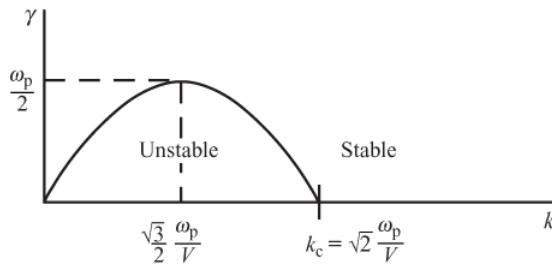


Fig. 8.3 A plot of the growth rate, γ , of the unstable solution as a function of the wave number, k , for two oppositely directed beams of equal density.

which can be rewritten in the form

$$\left(1 - \frac{\omega_p^2}{\omega^2}\right)(kV - \omega)^2 = \omega_b^2. \quad (8.1.37)$$

If the term on the right is small, one can assume that the solution is essentially the same as with no beam (i.e., $\omega = \omega_p$, and $k = \omega_p/V$) plus a small frequency shift $\omega = \omega_p + \delta\omega$. Substituting these approximations into the above equation and expanding to first order in $\delta\omega/\omega_p$ gives

$$\left[1 - \frac{\omega_p^2}{\omega_p^2} \left(1 - 2\frac{\delta\omega}{\omega_p}\right)\right] (\delta\omega)^2 = \omega_b^2. \quad (8.1.38)$$

This equation simplifies to $\delta\omega^3 = (1/2)\omega_b^2\omega_p$, which has three roots

$$\delta\omega = \left(\frac{1}{2}\right)^{1/3} \omega_b^{2/3} \omega_p^{1/3}, \quad \text{and} \quad \delta\omega = \left(\frac{1}{2}\right)^{1/3} \omega_b^{2/3} \omega_p^{1/3} \left[-\frac{1}{2} \pm i\frac{\sqrt{3}}{2}\right]. \quad (8.1.39)$$

The root with the positive imaginary part corresponds to an exponentially growing oscillation. It is easy to show that the frequency ω_r and growth rate γ of the unstable solution are given by

$$\omega_r = \omega_p \left[1 - \frac{1}{2} \left(\frac{1}{2}\right)^{1/3} \left(\frac{n_b}{n_0}\right)^{1/3}\right] \quad \text{and} \quad \gamma = \frac{\sqrt{3}}{2} \left(\frac{1}{2}\right)^{1/3} \left(\frac{n_b}{n_0}\right)^{1/3} \omega_p. \quad (8.1.40)$$

This beam-driven mode is shifted downward in frequency from the plasma frequency and grows at a rate proportional to the beam density to the one-third power, $\gamma \propto n_b^{1/3}$. Note that as the beam density goes to zero, this mode reduces to an undamped oscillation at the plasma frequency.

8.1.2.3 The Buneman instability

Another important instability, first investigated by Buneman [1959], consists of a cold electron beam streaming through a single species of positive ions initially at rest. Using equation (8.1.29), it is easy to show that the dispersion relation is given by

$$D(k, \omega) = 1 - \frac{\omega_{\text{pi}}^2}{\omega^2} - \frac{\omega_{\text{pe}}^2}{(kV - \omega)^2} = 0, \quad (8.1.41)$$

where ω_{pi} and ω_{pe} are the ion and electron plasma frequencies and $n_i = n_e = n_0$.

Note that since the ions are initially assumed to be at rest, the beam velocity associated with the ion term is zero. If one assumes that $\omega_{\text{pi}}^2 \ll \omega^2$ (which must be confirmed later), the above equation can be written in the approximate form,

$$kV - \omega = \pm \omega_{\text{pe}} \left[1 + \frac{\omega_{\text{pi}}^2}{2\omega^2} \right], \quad (8.1.42)$$

where the $(1 - \omega_{\text{pi}}^2/\omega^2)^{1/2}$ term has been expanded to first order in the small quantity $\omega_{\text{pi}}^2/\omega^2$. The minus sign gives a purely real solution for the frequency. For the plus sign, one can show by direct substitution that the frequency is complex and is given by

$$\omega = (\omega_{\text{pe}} \omega_{\text{pi}}^2 \cos \theta)^{1/3} e^{i\theta} \quad (8.1.43)$$

where θ is a parameter determined by k . Since this solution is complex, the corresponding mode is unstable whenever $\theta > 0$. By substituting $\omega = \omega_r + i\gamma$ into the above equation and separating the real and imaginary parts, it is easy to show that the maximum growth rate occurs at $\theta = \pi/3$. The growth rate and frequency at $\theta = \pi/3$ are given by

$$\gamma = \frac{\sqrt{3}}{2} \left(\frac{m_e}{2m_i} \right)^{1/3} \omega_{\text{pe}}, \quad (8.1.44)$$

and

$$\omega_r = \frac{1}{2} \left(\frac{m_e}{2m_i} \right)^{1/3} \omega_{\text{pe}}. \quad (8.1.45)$$

This very fast growing instability is called the Buneman mode and occurs whenever a cold electron beam streams through a background of cold ions. Later, when we consider beams with a finite temperature, we will show that the drift velocity must exceed the electron thermal speed before the instability can occur. Also, we will show that the finite temperature introduces another instability called the ion acoustic instability.

8.2 The Landau approach

As pointed out in Section 8.1, if the velocity distribution function is non-zero at a velocity equal to the phase velocity of the wave, the dispersion relation does not exist because of the singularity at $v_z = \omega/k$. This difficulty is inherent in the Vlasov approach, which does not properly treat the initial transient phase of a disturbance and, more importantly, assumes that a normal mode of the form $\exp(-i\omega t)$ exists for all times. In a classic paper, Landau [1946] resolved this difficulty by pointing out that it is necessary to consider the long-time response to a disturbance imposed at $t = 0$. This type of problem is called an initial value problem and is usually solved by using Laplace transforms. Since the solution involves some subtle points regarding the use of Laplace transforms, we start by giving a brief review of Laplace transforms.

8.2.1 Laplace transforms: a brief review

The Laplace transform of a function $f(t)$ is defined by

$$\tilde{f}(p) = \int_0^{\infty} f(t) e^{-pt} dt, \quad (8.2.1)$$

where p is a complex number, $p = \gamma - i\omega$. (Here we have defined the imaginary part of p as $-i\omega$ with ω real in order to maintain consistency with the Fourier analysis notation.) Since the integral diverges if the real part of p is negative, denoted by $\text{Re}\{p\} < 0$, the transform, $\tilde{f}(p)$, exists only if $\text{Re}\{p\} > 0$ and only for functions that grow less rapidly than an exponential. The inverse of a Laplace transform is given by the integral

$$f(t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \tilde{f}(p) e^{pt} dp, \quad (8.2.2)$$

where the integration path is along a line $\text{Re}\{p\} = \sigma$ that is located to the right of all the poles of $\tilde{f}(p)$, which are points where the function $\tilde{f}(p)$ goes to infinity. The required integration path is shown in Figure 8.4, where the “ \times ” symbols indicate the poles of $\tilde{f}(p)$.

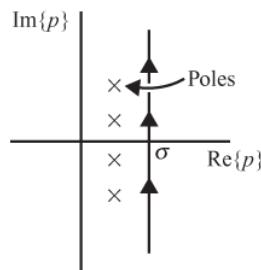


Fig. 8.4 The integration contour for the inverse Laplace transform must be along a line $\sigma = \text{Re}\{p\}$ that passes to the right of all poles of $\tilde{f}(p)$.

By integrating equation (8.2.1) by parts once, a useful relation can be established for the Laplace transform of the derivative of a function

$$\begin{aligned}\tilde{f}'(p) &= \int_0^\infty f'(t)e^{-pt} dt \\ &= [f(t)e^{-pt}]_0^\infty + p \int_0^\infty f(t)e^{-pt} dt \\ &= p\tilde{f}(p) - f(0).\end{aligned}\quad (8.2.3)$$

Note that the Laplace transform of the derivative of a function involves the initial value of the function, $f(0)$, evaluated at $t = 0$. By repeated use of the above equation, one can obtain the Laplace transform of higher order derivatives. For example, the Laplace transform of the second derivative of a function is given by

$$\tilde{f}''(p) = p[p\tilde{f}(p) - f(0)] - f'(0). \quad (8.2.4)$$

Laplace transforms can be used to obtain solutions of differential equations in much the same way as Fourier transforms. However, Laplace transforms have the advantage that the initial conditions are automatically included in the solution.

To illustrate the use of Laplace transforms for solving a differential equation, we solve the problem of a damped harmonic oscillator subject to a specific set of initial conditions at $t = 0$. The differential equation for a damped harmonic oscillator is given by

$$f'' + 2\gamma_0 f' + \omega_0^2 f = 0, \quad (8.2.5)$$

where γ_0 is the damping constant and ω_0 is the characteristic frequency. The Laplace transform of the above equation is

$$\tilde{f}'' + 2\gamma_0 \tilde{f}' + \omega_0^2 \tilde{f} = 0. \quad (8.2.6)$$

Using equations (8.2.3) and (8.2.4), we obtain

$$p^2 \tilde{f} - pf(0) - f'(0) + 2\gamma_0 p \tilde{f} - 2\gamma_0 f(0) + \omega_0^2 \tilde{f} = 0. \quad (8.2.7)$$

This purely algebraic equation can then be solved for \tilde{f} , which gives

$$\tilde{f}(p) = \frac{f'(0) + (p + 2\gamma_0)f(0)}{p^2 + 2\gamma_0 p + \omega_0^2}. \quad (8.2.8)$$

To obtain $f(t)$, we must evaluate the inverse Laplace transform of this function. One way to obtain the inverse Laplace transform is by simply looking up the inverse in a table. However, this cannot always be done and for our purposes it is instructive to compute the inverse directly by evaluating the integral in equation (8.2.2). Complex integrals of this type are usually evaluated using the Cauchy residue theorem [see Arfken, 1970]. This theorem allows one to compute the contour integral of a function $g(z)$ that is analytic except for

a finite number of poles. An analytic function is a function that has a continuous unique derivative in some region R of the complex plane. Any function that can be expanded as a power series is an analytic function. If the function contains a first-order pole z_0 of the form $g(z) = G(z)/(z - z_0)$ within a contour C , then by shrinking the contour to a small circle around z_0 one can show that

$$\int_C g(z) dz = \int_C \frac{G(z)}{z - z_0} dz = 2\pi i \operatorname{Res}[g(z), z_0] = 2\pi i G(z_0), \quad (8.2.9)$$

where the quantity $\operatorname{Res}[g(z), z_0]$ is called the residue of the function $g(z)$ evaluated at z_0 . For a first-order pole, the residue is simply the function $G(z)$. If the function contains several first-order poles, then the integral is simply $2\pi i$ times the sum of the residues within the contour. For a pole of order n , such that $g(z) = G(z)/(z - z_0)^n$, the residue is given by

$$\operatorname{Res}[g(z), z_0] = \frac{1}{(n-1)!} \left. \frac{d^{n-1}}{dz^{n-1}} G(z) \right|_{z=z_0}. \quad (8.2.10)$$

To evaluate the inverse Laplace transform using the residue theorem, the line of integration, $\operatorname{Re}\{p\} = \sigma$, must be distorted into a closed contour. To assure that the additional integration path produces no contribution to the integral, the line of integration must be closed at infinity in the left half of the complex p -plane, $\operatorname{Re}\{p\} < 0$, as shown in Figure 8.5. The integral diverges if the contour is closed at infinity in the right half-plane. Since the Laplace transform is only defined for $\operatorname{Re}\{p\} > 0$, the function $\tilde{f}(p)$ must be extended into the left half of the complex p -plane by analytic continuation [see Flanigan, 1972]. In the case of functions involving the ratio of simple polynomials, the analytic continuation is the same function, with unchanged coefficients for the polynomials. In this case, the analytic continuation into the left half-plane is trivial; it is the same function. Later we will encounter a case where the analytic continuation must be treated more delicately.

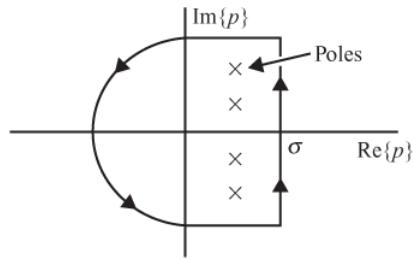


Fig. 8.5

To assure convergence, the integral in the inverse Laplace transform must be closed at infinity in the left half of the complex p -plane.

To use the residue theorem, we must determine the location of the poles of $\tilde{f}(p)$. As can be seen by inspection of equation (8.2.8), the poles are located at the points where the denominator is zero, which are at

$$p^\pm = -\gamma_0 \pm i \sqrt{\omega_0^2 - \gamma_0^2}. \quad (8.2.11)$$

The inverse Laplace transform is then given by

$$f(t) = \frac{1}{2\pi i} \int_C \left[\frac{f'(0) + (p + 2\gamma_0)f(0)}{(p - p_+)(p - p_-)} \right] e^{pt} dp, \quad (8.2.12)$$

where the contour C now passes around the poles at p_+ and p_- . Evaluating the appropriate residues, it is easy to show that

$$f(t) = \frac{f'(0) + (p_- + 2\gamma_0)f(0)}{p_- - p_+} e^{p_- t} + \frac{f'(0) + (p_+ + 2\gamma_0)f(0)}{p_+ - p_-} e^{p_+ t}, \quad (8.2.13)$$

which after substituting for p_+ and p_- simplifies to

$$f(t) = \frac{f'(0) + \gamma_0 f(0)}{\omega} e^{-\gamma_0 t} \sin \omega t + f(0) e^{-\gamma_0 t} \cos \omega t, \quad (8.2.14)$$

where $\omega = \sqrt{\omega_0^2 - \gamma_0^2}$. Note that the roots of the denominator of $\tilde{f}(p)$ completely determine the oscillation frequency and damping.

8.2.2 The dispersion relation for a plasma of hot electrons and immobile ions

As an initial application of the Landau method, we start by considering the small amplitude waves that can occur in a plasma of hot electrons and immobile ions. The relevant equations are (8.1.4) and (8.1.5). To solve these equations for an arbitrary initial condition, we perform a Fourier transform in space (with the \mathbf{k} vector along the z axis) and a Laplace transform in time. After dropping the subscript 1 on the first-order terms, the transforms of these equations are

$$p\tilde{f} - f(0) + ikv_z\tilde{f} + i\frac{e}{m} k\tilde{\Phi} \frac{\partial f_0}{\partial v_z} = 0, \quad (8.2.15)$$

and

$$k^2\tilde{\Phi} = -\frac{e}{\epsilon_0} \int_{-\infty}^{\infty} \tilde{f} d^3v, \quad (8.2.16)$$

where $f(0)$ is the first-order distribution function evaluated at $t = 0$. Solving equation (8.2.15) for \tilde{f} gives

$$\tilde{f} = \frac{-i\frac{e}{m} k\tilde{\Phi} \frac{\partial f_0}{\partial v_z} + f(0)}{p + ikv_z}. \quad (8.2.17)$$

Substituting this expression into equation (8.2.16) and introducing the reduced distribution functions, $F_0(v_z) = (1/n_0) \int f_0 dv_x dv_y$ and $F(0) = (1/n_0) \times \int f(0) dv_x dv_y$, we obtain the equation

$$k^2 \tilde{\Phi} = -\frac{en_0}{\epsilon_0} \int_{-\infty}^{\infty} \frac{F(0)}{p + ikv_z} dv_z + i\omega_p^2 k \tilde{\Phi} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{p + ikv_z} dv_z, \quad (8.2.18)$$

where $F(0)$ is the reduced first-order distribution function evaluated at $t = 0$. The above equation can then be solved for $\tilde{\Phi}$, and written in the form

$$\tilde{\Phi}(k, p) = \frac{N(k, p)}{D(k, p)}, \quad (8.2.19)$$

where the numerator is given by

$$N(k, p) = i \frac{en_0}{\epsilon_0 k^3} \int_{-\infty}^{\infty} \frac{F(0)}{v_z - ip/k} dv_z, \quad (8.2.20)$$

and the denominator is given by

$$D(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{v_z - ip/k} dv_z. \quad (8.2.21)$$

Note that the denominator can be obtained by simply changing ω to ip in the dispersion relation obtained from the Fourier analysis approach, i.e., equation (8.1.13).

To determine the temporal behavior of the electrostatic potential we must carry out the inverse Laplace transform

$$\Phi(k, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \tilde{\Phi}(k, p) e^{pt} dp, \quad (8.2.22)$$

where the integration is along the line $\text{Re}\{p\} = \sigma$. As discussed earlier, this line must be to the right of all poles of $\tilde{\Phi}(k, p)$.

Following the procedure discussed in Section 8.2.1, we evaluate the inverse Laplace transform by closing the contour of integration in the left half-plane, $\text{Re}\{p\} < 0$, and distorting the integration contour so that the integral can be evaluated by integrating around the poles of $\tilde{\Phi}(k, p)$ using the residue theorem. These steps are illustrated in Figure 8.6. To use the residue theorem, there are two requirements. First, the function $\tilde{\Phi}(k, p)$ must be analytic except for an isolated number of poles. This means that we cannot consider velocity distribution functions that have discontinuities, such as step functions and various other pathological functions. To meet this requirement, we restrict the analysis to distribution functions that are analytic everywhere, that is, can be represented by a power series, $F_0(v_z) = \sum_n a_n v_z^n$. Second, the function $\tilde{\Phi}(k, p)$ must be analytically continued from the right half of the complex p -plane into the left half-plane. Unfortunately, this analytic continuation is not straightforward, since the integrals in $N(k, p)$ and $D(k, p)$ have a discontinuity along the line $\gamma = \text{Re}\{p\} = 0$, which separates the right-hand and left-hand halves of the complex p -plane. This difficulty is closely related to the problem that was encountered in the Vlasov approach, namely the singularity at $v_z = \omega/k$.

The resolution of the difficulty at $v_z = \omega/k$ was provided by Landau, who carried out

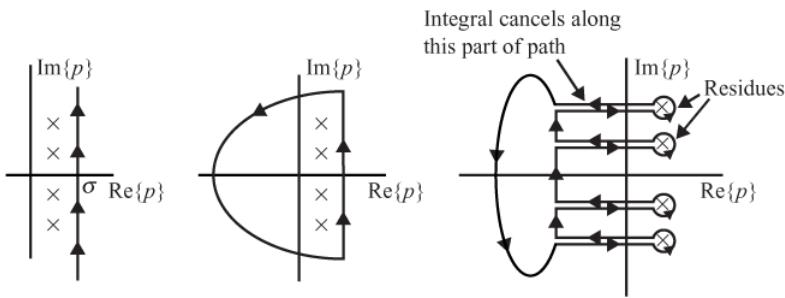


Fig. 8.6

The final stage in evaluating the inverse Laplace transform involves distorting the integration contour around the poles and arranging for the integral to cancel around the rest of the contour. The resulting integrals around the poles are called the residues.

a careful analytic continuation of the functions $N(k, p)$ and $D(k, p)$. The procedure is essentially the same for the numerator and denominator, so we only discuss the denominator. The requirement for an analytic continuation is that the function must vary smoothly across the boundary between the two regions. From the Laplace transform analysis given above, equation (8.2.21) is a valid representation of $D(k, p)$ in the right half-plane, $\text{Re}\{p\} > 0$. Since the sign of k determines the sign of ip/k , we must consider two cases, k positive and k negative. First, we consider the case where k is positive. Even though the integral in equation (8.2.21) is along the real velocity axis, it is useful to visualize the integration path in the complex v_z plane, as shown in Figure 8.7. For $\text{Re}\{p\} > 0$ and $k > 0$, the pole at $v_z = ip/k$ lies above the real v_z axis. As the real part of p changes from positive to negative, the pole moves downward across the real v_z axis. To avoid a discontinuity when the pole crosses the real v_z axis, Landau simply distorted the integration contour so that it always stays below the pole, as shown in Figure 8.8. It is then a simple matter to define a function $D(k, p)$ that is valid for any value of $\text{Re}\{p\}$.

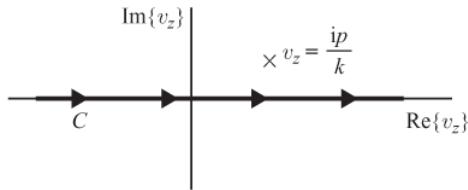


Fig. 8.7

For $\text{Re}\{p\}$ positive the integration contour in $D(k, p)$ is along the $\text{Re}\{v_z\}$ axis.

This function is given by

$$D(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_C \frac{\partial F_0 / \partial v_z}{v_z - ip/k} dv_z. \quad (8.2.23)$$

where the contour C passes below the pole at $v_z = ip/k$. Note that the only difference in the two cases is the shape of the integration contour. For $\text{Re}\{p\} > 0$, the integration

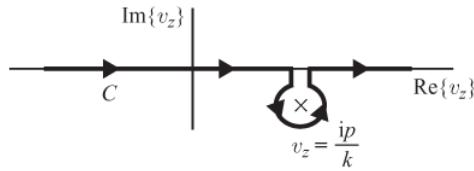


Fig. 8.8 To continue $D(k, p)$ analytically into the left half of the complex p -plane the integration contour must be distorted such that it passes below the pole at $v_z = ip/k$. This diagram is for $k > 0$.

contour extends along the real v_z axis from minus infinity to plus infinity. For $\text{Re}\{p\} < 0$ the integration contour includes an additional loop around the pole at $v_z = ip/k$.

Another way to view the above analytical continuation is to define two functions $D_+(k, p)$ and $D_-(k, p)$ that are valid for $\text{Re}\{p\} > 0$ and $\text{Re}\{p\} < 0$. These functions are

$$D_+(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{v_z - ip/k} dv_z, \quad \text{Re}\{p\} > 0 \quad (8.2.24)$$

and

$$\begin{aligned} D_-(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{v_z - ip/k} dv_z \\ - 2\pi i \frac{k}{|k|} \frac{\omega_p^2}{k^2} \frac{\partial F_0}{\partial v_z} \Big|_{v_z=ip/k}, \quad \text{Re}\{p\} < 0. \end{aligned} \quad (8.2.25)$$

The $2\pi i$ term in equation (8.2.25) above comes from evaluating the residue around the pole at $v_z = ip/k$. This term corrects for the discontinuity as the pole crosses the real v_z axis. Note that the integrals in both of the above two equations are along the real v_z axis from $-\infty$ to $+\infty$. The $k/|k|$ term in the residue takes care of the case where k is negative, which has the effect of reversing the sense (hence sign) of the contour around the pole. An essentially identical procedure is used to obtain the analytic continuation of $N(k, p)$.

Having established the analytic continuation of $N_+(k, p)$ and $D_+(k, p)$, the temporal variation of the electrostatic potential can be determined by computing the inverse Laplace transform

$$\Phi(k, t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \tilde{\Phi}(k, p) e^{pt} dt. \quad (8.2.26)$$

The inverse transform is easily evaluated using the residue theorem, which gives a solution of the form

$$\Phi(k, t) = \sum_i e^{p_i t} \text{Res}[\tilde{\Phi}(k, p), p_i], \quad (8.2.27)$$

where p_i is the i th pole of $\tilde{\Phi}(k, p)$. Since $F(0)$ was assumed to be an analytic function of v_z ,

the numerator of $\tilde{\Phi}(k, p)$ does not have any singularities. Hence, the only poles of $\tilde{\Phi}(k, p)$ are those for which $D(k, p) = 0$.

Because the p_i are in general complex, the time behavior of the potential (hence the electric field) is in general oscillatory, with exponential damping (growth) if the real part of p_i is negative (positive). As t becomes large, the pole with the largest value for $\text{Re}\{p_i\}$ dominates the oscillatory behavior of the electric field.

Since the roots of $D(k, p) = 0$ give the frequency and growth rate of the electrostatic potential, the equation $D(k, p) = 0$ plays the same role as the dispersion relation in the Fourier analysis approach. However, the roots, $p_i = \gamma_i - i\omega_i$, are now complex.

8.2.3 The Cauchy velocity distribution function

As a simple example that illustrates the above solution, consider a plasma consisting of electrons with a velocity distribution function given by

$$F_0(v_z) = \frac{C}{\pi} \frac{1}{C^2 + v_z^2}, \quad (8.2.28)$$

and an equal density of immobile ions. This distribution function, which is called the Cauchy distribution function, has the advantage of being relatively easy to analyze and has a shape similar to a Maxwellian. The easiest way to evaluate the dispersion relation is by integrating equation (8.2.23) by parts once, which gives

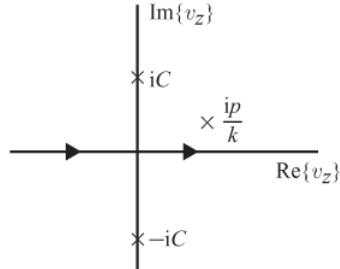


Fig. 8.9 The location of the poles in the integral for $D(k, p)$ for the Cauchy distribution function. This diagram assumes that $\text{Re}\{p\}$ is positive and that $k > 0$.

$$D(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_C \frac{F_0(v_z)}{(v_z - ip/k)^2} dv_z = 0. \quad (8.2.29)$$

After substituting $F_0(v_z)$ from equation (8.2.28) into the above equation, the dispersion relation can be written

$$D(k, p) = 1 - \frac{\omega_p^2 C}{k^2 \pi} \int_C \frac{dv_z}{(v_z - iC)(v_z + iC)(v_z - ip/k)^2} = 0. \quad (8.2.30)$$

The integrand in the above equation has three poles, at $v_z = iC$, $-iC$, and ip/k . For purposes of computing the integral, we start by assuming that $\text{Re}\{p\}$ is positive. Two cases must be

considered, $k > 0$ and $k < 0$. For $k > 0$, the poles are located as shown in Figure 8.9. The integral is easily evaluated using the residue theorem. In this case it is convenient to close the contour in the lower half of the complex v_z plane for $k > 0$, and in the upper half of the complex v_z plane for $k < 0$. This procedure minimizes the number of residues that must be considered. After evaluating the residues, it is easy to show that the dispersion relation can be written as a single equation that is valid for both $k > 0$ and $k < 0$

$$D(k, p) = 1 + \frac{\omega_p^2}{(p + C|k|)^2} = 0. \quad (8.2.31)$$

At this point the dispersion relation is only valid in the right half of the complex p -plane, since we have assumed that $\text{Re}\{p\} > 0$. However, the analytic continuation to the lower half-plane is trivial, since the function is a polynomial in p and is therefore continuous throughout the complex p -plane. Note that since a direct analytical continuation can be made, it is not necessary to consider the alternative definitions for $D(k, p)$ given by equation (8.2.23), or equations (8.2.24) and (8.2.25).

The roots to the above dispersion relation are easily shown to be

$$p + C|k| = \pm i\omega_p. \quad (8.2.32)$$

Since $p = \gamma - i\omega$, the growth rate and frequency are $\gamma = -C|k|$ and $\omega = \pm\omega_p$. The temporal dependence of the electrostatic potential is then given by

$$\tilde{\Phi}(k, t) = \tilde{\Phi}_0 e^{\pm i\omega_p t} e^{-C|k|t}, \quad (8.2.33)$$

where $\tilde{\Phi}_0$ represents the contribution obtained from the initial value term. Note that the growth rate, $\gamma = -C|k|$, is negative. This means that the fields are exponentially damped. This damping is called Landau damping, after Landau [1946] who first identified this collisionless damping process.

8.2.4 The weak growth rate approximation

To gain further insight into the nature of Landau damping, we next develop an approximate solution of the dispersion relation that is valid whenever the magnitude of the growth rate is small compared to the frequency, $|\gamma| \ll |\omega|$. This approximation is obtained by expanding the dispersion relation in a Taylor series around the point $p = -i\omega$, where $p = \gamma - i\omega$. To first order in the small parameter γ , the Taylor series expansion of $D(k, p)$ is

$$D(k, p) = D(k, -i\omega) + i \left(\frac{\partial D(k, -i\omega)}{\partial \omega} \right) \gamma = 0. \quad (8.2.34)$$

The first term, $D(k, -i\omega)$, includes both real and imaginary parts, both of which must be evaluated in the limit $\gamma \rightarrow 0$. Thus, for the first term we write

$$D(k, -i\omega) = D_r(k, -i\omega) + iD_i(k, -i\omega). \quad (8.2.35)$$

The second term, $(\partial D / \partial \omega) \gamma$, also includes both real and imaginary parts. However, we only keep the real part, since for small $|\gamma|$ the imaginary part gives a negligible correction to the real part of equation (8.2.34). Combining the above two equations and setting the real and imaginary parts to zero then gives two equations,

$$D_r = 0, \quad (8.2.36)$$

and

$$\gamma = \frac{-D_i}{\partial D_r / \partial \omega}. \quad (8.2.37)$$

Since we are taking the limit as γ approaches zero, to evaluate D_i and D_r we must carefully consider the integral in $D(k, -i\omega)$, which after substituting $p = \gamma - i\omega$ can be written



Fig. 8.10 In the limit $\epsilon \rightarrow 0$ the integration contour in the Plemelj relation goes half-way around the pole at x_0 .

$$D(k, -i\omega) = 1 - \lim_{\gamma \leftarrow 0} \frac{\omega_p^2}{k^2} \int_C \frac{\partial F_0 / \partial v_z}{v_z - (\omega/k + i\gamma/k)} dv_z. \quad (8.2.38)$$

To evaluate this integral we use the Plemelj relation [Dennery and Krzywicki, 1967]. The Plemelj relation is given by

$$\lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} \frac{f(x)}{x - (x_0 \pm i\epsilon)} dx = P \int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx \pm i\pi f(x_0), \quad (8.2.39)$$

where ϵ is a positive quantity and the P refers to the principal value integral, which is defined by

$$P \int_{-\infty}^{\infty} \dots dx = \lim_{\delta \leftarrow 0} \left[\int_{-\infty}^{x_0 - \delta} \dots dx + \int_{x_0 + \delta}^{\infty} \dots dx \right]. \quad (8.2.40)$$

Note that the usual $2\pi i$ term in the residue around the pole is replaced by $i\pi$. The change from $2\pi i$ to $i\pi$ occurs because in the limit $\epsilon \rightarrow 0$ the integration contour only goes half-way around the pole, as shown in Figure 8.10. After using the Plemelj relation, D_r and D_i become

$$D_r = 1 - \frac{\omega_p^2}{k^2} P \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{v_z - \omega/k} dv_z, \quad (8.2.41)$$

and

$$D_i = -\pi \frac{k}{|k|} \frac{\omega_p^2}{k^2} \left. \frac{\partial F_0}{\partial v_z} \right|_{v_z=\omega/k}. \quad (8.2.42)$$

The $k/|k|$ term in D_i takes into account the change in sign that occurs when $k > 0$ and $k < 0$. As in the Cauchy distribution, the sign of k determines whether the pole approaches the real v_z axis from above ($k > 0$) or from below ($k < 0$), which in turn determines the \pm sign in the Plemelj relation. The real part of D is essentially the same as was obtained from the Fourier analysis approach, except the principal value integral now provides a mathematical procedure for avoiding the divergence in the integral at $v_z = \omega/k$. Substituting equation (8.2.42) into equation (8.2.37), we obtain the following expression for the growth rate

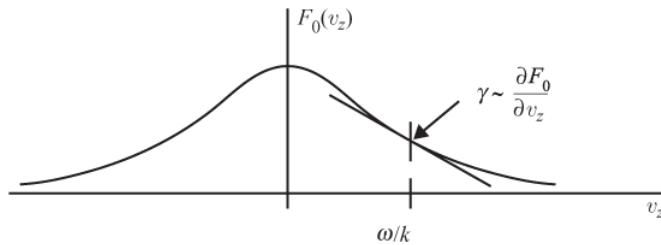


Fig. 8.11

In the weak growth rate approximation, the growth rate, γ , is proportional to the slope of the distribution function evaluated at the phase velocity, ω/k , of the wave.

$$\gamma = \pi \frac{k}{|k|} \frac{\omega_p^2}{k^2 \partial D_r / \partial \omega} \left. \frac{\partial F_0}{\partial v_z} \right|_{v_z=\omega/k}. \quad (8.2.43)$$

This result is called the weak growth rate approximation and is valid whenever $|\gamma| \ll |\omega|$. The weak growth rate approximation shows that the growth rate is proportional to the slope of the reduced distribution evaluated at the phase velocity of the wave (see Figure 8.11).

For Langmuir waves, we can carry out an explicit evaluation of the $\partial D_r / \partial \omega$ term in equation (8.2.43). From the real part of the dispersion relation, it is easy to show, using the same procedure as in the Fourier analysis approach, that in the high-phase velocity limit the real part of the dispersion relation is given by

$$D_r = 1 - \frac{\omega_p^2}{\omega^2} \left(1 + 3 \frac{k^2}{\omega^2} \langle v_z^2 \rangle \right). \quad (8.2.44)$$

To lowest order k , $D_r = 1 - \omega_p^2 / \omega^2$, so $\partial D_r / \partial \omega = 2\omega_p^2 / \omega^3$ which, when evaluated at $\omega = \omega_p$, gives

$$\gamma = \frac{\pi}{2} \frac{k}{|k|} \frac{\omega_p^3}{k^2} \left. \frac{\partial F_0}{\partial v_z} \right|_{v_z=\omega/k}. \quad (8.2.45)$$

If the reduced distribution function has a single maximum at $v_z = 0$, such as for a Maxwellian or a Cauchy distribution, one can see from the above equation that the growth rate is always

negative (i.e., indicating damping), independent of the direction of propagation. Note that when k is positive, ω/k is positive and $\partial F_0/\partial v_z$ is negative, which makes γ negative. When k is negative, ω/k is negative, and $\partial F_0/\partial v_z$ is positive, which again makes γ negative.

For a Maxwellian distribution, it is easy to show that the frequency and growth rate are given, respectively, by

$$\omega^2 = \omega_p^2 + 3\left(\frac{\kappa T}{m}\right) k^2, \quad (8.2.46)$$

and

$$\gamma = -\sqrt{\frac{\pi}{8}} \frac{\omega_p}{|k\lambda_D|^3} \exp\left[-\frac{1}{2(k\lambda_D)^2} - \frac{3}{2}\right], \quad (8.2.47)$$

where λ_D is the Debye length. The frequency is identical to the result obtained in the Fourier analysis approach (i.e., the Bohm–Gross dispersion relation). The growth rate is always negative which indicates damping, very similar to the situation with the Cauchy distribution. Note that the $\exp[-3/2]$ factor comes from the second term on the right side of equation (8.2.46). Because of the strong exponential dependence of the Maxwellian, it is important to keep this higher order term. As can be seen, the damping is a strong function of $k\lambda_D$. For long wavelengths, much greater than the Debye length, the damping is very weak. As the wavelength decreases, the damping increases rapidly, becoming very strong as the wavelength approaches the Debye length. Oscillatory solutions do not exist for $k\lambda_D \gtrsim 1$. Physically, the damping becomes large as the phase velocity of the wave moves into the region where the slope of the distribution function is large. A plot of the frequency, ω , and damping rate, $-\gamma$, is shown in Figure 8.12. The dashed lines are from equations (8.2.46) and (8.2.47), and the solid lines are from an exact numerical solution of the dispersion relation. As can be seen, equations (8.2.46) and (8.2.47) give a good fit for $k\lambda_D \ll 1$, but fail as $k\lambda_D \rightarrow 1$.

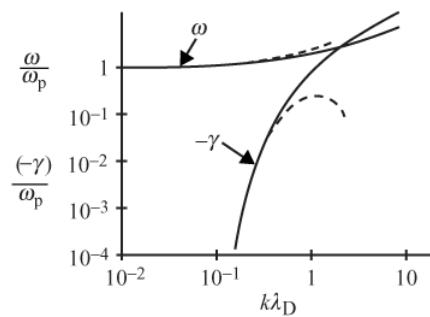


Fig. 8.12 A comparison of an approximate analytical solution to the Langmuir wave dispersion relation (dashed lines) with the results of an exact numerical computation (solid lines).

8.2.5 The physical origin of Landau damping

The theoretical discovery by Landau [1946] of a wave damping mechanism that does not depend on collisions was not only surprising, but left many unanswered questions regarding the physical processes involved. That the damping is real, and not just an abstract result of a contour integration, was first shown in a laboratory experiment by Malmberg and Wharton [1966] and has since been confirmed in a wide variety of laboratory and space plasmas. From our previous analyses, which showed that the growth rate is proportional to the slope of the zero-order distribution function evaluated at the phase velocity of the wave, it is clear that Landau damping is somehow closely associated with particles moving at velocities near the phase velocity of the wave, the so-called resonant particles. To understand the physical origin of the damping, it is useful to transform into a frame of reference that is moving at the phase velocity of the wave. The advantage of this frame of reference is that for a sufficiently small growth rate the electrostatic potential of the wave is nearly time stationary. The total energy of the particles, $W = (1/2)mv_z^2 + q\Phi$, is then conserved, at least to a first approximation. The particle velocities are then given by

$$v_z = \pm \sqrt{\frac{2}{m}(W - q\Phi)}, \quad (8.2.48)$$

where the (+) sign is for particles moving in the $+z$ direction, and the (-) sign is for particles moving in the $-z$ direction. For an electrostatic potential of the form $\Phi(z) = \Phi_0 \cos(kz)$, the trajectories in (z, v_z) phase space are as shown in Figure 8.13. The slightly darker lines in the phase-space plot that separate trapped from untrapped particles are called separatrices. The total energy of particles on the separatrices is given by $W_0 = q\Phi_0$, where Φ_0 is the amplitude of the electrostatic potential. Particles with total energies $W < W_0$ are trapped and particles with total energies $W > W_0$ are untrapped. The velocity distribution of particles on these trajectories is determined by the zero-order distribution function, $F_0(v_z)$. Once the initial particle distribution function is determined, trapped particles, such as A and B, follow the elliptical shaped trajectories inside the separatrices, and untrapped particles, such as A' and B', stream freely along the trajectories outside the separatrices. The oscillation frequency of the trapped particles depends on how deep the particles are in the potential well. Particles trapped near the bottom of the potential well, where the potential is parabolic, oscillate at a frequency called the bounce frequency, which is easily shown to be

$$\omega_b = k \sqrt{\frac{q\Phi_0}{m}}. \quad (8.2.49)$$

Due to the increasing width of the potential well at higher energies, the oscillation frequency decreases with increasing total energy, ultimately going to zero for particles at the top of the well. Since the oscillation frequency decreases with increasing energy, as the particles move along the elliptical shaped trajectories, they gradually get out of phase. This process is called phase mixing and is illustrated in Figure 8.14, which shows the relative locations of particles trapped in the potential well at three different times. Although the

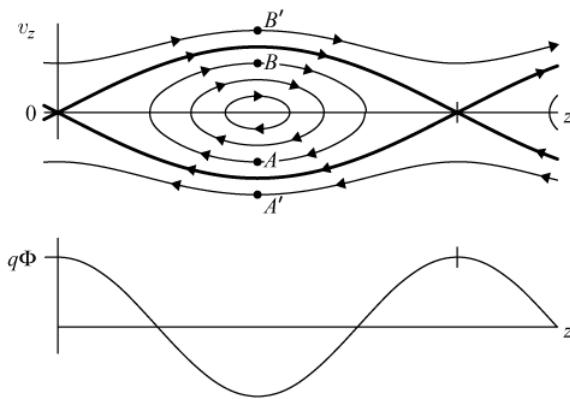


Fig. 8.13 Phase-space trajectories in a frame of reference (z, v_z) moving at the phase velocity for a sinusoidal electrostatic potential $\Phi(z) = \Phi_0 \cos(kz)$.

trapped particle motions may be highly correlated initially, they eventually become well mixed throughout the trapping region. As we shall see, it is this phase mixing that ultimately accounts for the dissipation that is implicit in Landau damping.

Since the equations that were used to derive Landau damping were linearized, it is clear that Landau damping is an inherently linear process. However, there is also a non-linear aspect to Landau damping, and it is important to recognize the difference between the two processes. That non-linear effects eventually become involved is easily demonstrated. From the derivation of equation (8.1.4), one can see that the key requirement in the linearization process is that

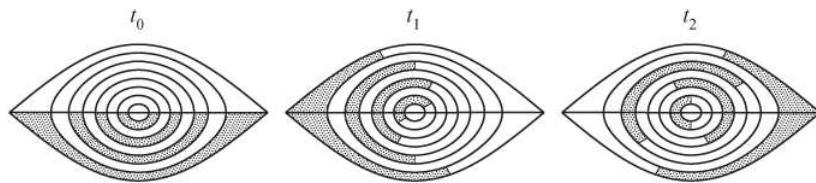


Fig. 8.14 The relative phase-space locations of trapped particles at three successive phases of the bounce cycle.

$$|\partial f / \partial v_z| \ll |\partial f_0 / \partial v_z|. \quad (8.2.50)$$

Since the zero-order term, $\partial f_0 / \partial v_z$, is independent of time, we focus our attention on the first-order term, $\partial f / \partial v_z$. From equation (8.2.17), one can see that there is a term in $\partial \tilde{f} / \partial v_z$ of the form

$$\frac{\partial \tilde{f}}{\partial v_z} = -\frac{ikf(0)}{(p + ikv_z)^2}. \quad (8.2.51)$$

Using a table of inverse Laplace transforms, one can show that this term has a temporal dependence given by

$$\frac{\partial f}{\partial v_z} = -ikf(0)te^{-ikv_z t}. \quad (8.2.52)$$

Since the magnitude of this term increases linearly with time, it is clear that the inequality in equation (8.2.50) is eventually violated, no matter how small the initial perturbation, $f(0)$.

From the above discussion it is clear that there are two time scales that must be considered in Landau damping, an early phase in which linear analysis holds, and a later phase in which non-linear effects dominate. To distinguish the physical processes that occur during these two phases, we return to a discussion of particles trapped in the potential well of the wave. For sufficiently small wave amplitudes, equation (8.2.49) shows that the bounce period of the trapped particles is very long, much longer than the oscillation period of the electrostatic wave. Therefore, there is an initial phase during which particle trapping effects are unimportant, since the particles have not had time to complete even a small fraction of a bounce cycle. It is during this initial phase that Landau's linear analysis applies. The duration of the linear phase is bounded by

$$0 < t \ll \omega_b^{-1} = \frac{1}{k} \sqrt{\frac{m}{q\Phi_0}}. \quad (8.2.53)$$

The easiest way to understand the damping of the wave during the linear phase is to investigate the kinetic energy of the particles. Since the total energy of the system is conserved, if the kinetic energy of the particles increases, then the electrostatic energy of the wave must decrease. Whether the kinetic energy of the particles increases or decreases in response to the presence of the wave depends on the initial distribution of particle velocities, as characterized by the zero-order particle distribution function, $F_0(v_z)$. If the slope of the zero-order distribution function, $\partial F_0(v_z)/\partial v_z$, is negative (at $\omega/k > 0$), then the number of particles initially moving slower than the phase velocity (for example, particles A and A' in Figure 8.13) is greater than the number of particles initially moving faster than the phase velocity (for example, particles B and B' in Figure 8.13). As time advances, these particles proceed to move in the direction of the arrows. From the direction of motion it is evident that the kinetic energy in the plasma rest frame, which is proportional to $(\omega/k + v_z)^2$, increases for particles A and A' and decreases for particles B and B' . Since there are more particles of type A and A' than there are of type B and B' , the total kinetic energy contributed by these particles must increase. Note that the increase of the total kinetic energy in the plasma rest frame is a consequence of the quadratic dependence of the kinetic energy on v_z , i.e., $(\omega/k + v_z)^2$. Also, note that since the particles have only moved a small fraction of a bounce cycle, the issue of whether the particles are trapped or not is irrelevant during the initial linear phase. By carrying out a summation over all of the resonant particles, Nicholson [1983] and others have shown that the rate of change of the kinetic energy is given by

$$\frac{d}{dt} \left\langle \frac{1}{2} n_0 m \left(\frac{\omega}{k} + v_z \right)^2 \right\rangle = -\pi \left(\frac{\epsilon_0 E_0^2}{2} \right) \frac{\omega_p^3}{k^2} F'_0 \left(\frac{\omega}{k} \right), \quad (8.2.54)$$

where E_0 is the electric field amplitude. As one can see, if the zero-order distribution function, $F_0(\omega/k)$, is negative, the kinetic energy of the particles increases. The wave energy must then decrease, which implies damping. It can be shown that the above equation gives a damping rate that is in exact agreement with the damping rate obtained using the weak growth rate approximation, i.e., equation (8.2.45).

As the linear phase of the Landau damping proceeds, a point is eventually reached where inequality (8.2.53) is no longer satisfied. The bounce motion of the trapped particles then becomes important [O'Neil, 1965]. Since the trapped particle motions involve turning points, the dynamical motions are inherently non-linear. What happens during the non-linear phase is difficult to determine analytically and is usually investigated by solving numerically the fully non-linear equations. From such numerical computations, it is possible to give a relatively simple interpretation of the physical processes involved. During the initial linear phase, the wave gives energy to the resonant particles, both trapped and untrapped. As time progresses into the non-linear phase, the trapped particles start to give some of the energy back to the wave. However, not all of the initial energy can be returned, because some has been lost to the untrapped particles. Also, the trapped particles have started to phase mix, which diminishes the energy return. The result is that the wave amplitude decreases less rapidly than predicted by the linear damping rate and starts to level off, as shown in Figure 8.15, or may even increase [Armstrong, 1967]. Eventually, after many bounces, the trapped particles become thoroughly phase mixed. At this point, the wave damping ceases. Although the trapped particles in this final state may appear to be randomly distributed, they are not. In principle, the information imprinted by the initial state is still contained in their relative phases. However, if there is even the smallest amount of dissipation, such as due to collisions, the phase mixed distribution is eventually converted to a random thermalized distribution.

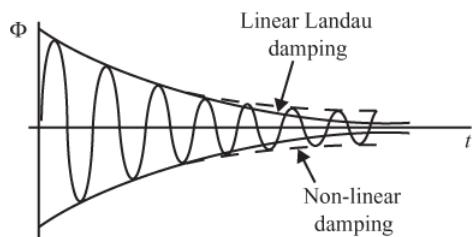


Fig. 8.15

The non-linear effects of particle trapping tend to increase the wave amplitude relative to the predictions of linear Landau damping.

8.3 The plasma dispersion function

Next we discuss the dispersion relation for a Maxwellian distribution function, which in normalized one-dimensional form is given by

$$F_0(v_z) = \left(\frac{m}{2\pi kT} \right)^{1/2} \exp \left[-\frac{mv_z^2}{2kT} \right]. \quad (8.3.1)$$

Unfortunately, for a Maxwellian distribution function, the integral in the dispersion relation (8.2.23) cannot be evaluated using the residue theorem. The reason is that when the integration contour is closed at infinity the integral diverges. Nevertheless, approximations can be carried out in certain limits, and it is useful to develop the tools necessary to carry out these approximations, since we will have use for these tools later. If we make a change of variables to

$$z = \sqrt{\frac{m}{2kT}} v_z \quad \text{and} \quad \zeta = \sqrt{\frac{m}{2kT}} \left(\frac{ip}{k} \right), \quad (8.3.2)$$

it is easy to show that the dispersion relation (8.2.23) can be rewritten in the form

$$D(k, p) = 1 + \frac{1}{(k\lambda_D)^2} \frac{1}{\sqrt{\pi}} \int_C \frac{z e^{-z^2}}{z - \zeta} dz = 0, \quad (8.3.3)$$

where we have made use of the fact that $\sqrt{kT/m} = \omega_p \lambda_D$. The integral in the above expression can be rewritten in a standard form by the following algebraic reorganization,

$$\begin{aligned} \frac{1}{\sqrt{\pi}} \int_C \frac{z e^{-z^2}}{z - \zeta} dz &= \frac{1}{\sqrt{\pi}} \int_C \left(1 + \frac{\zeta}{z - \zeta} \right) e^{-z^2} dz \\ &= 1 + \zeta Z(\zeta), \end{aligned} \quad (8.3.4)$$

where we have used $\int \exp[-z^2] dz = \sqrt{\pi}$, and the function

$$Z(\zeta) = \frac{1}{\sqrt{\pi}} \int_C \frac{e^{-z^2}}{z - \zeta} dz \quad (8.3.5)$$

is known as the plasma dispersion function [Fried and Conte, 1961]. The contour C is understood to be along the real z axis, passing under the pole at $z = \zeta$. Using the plasma dispersion function, the dispersion relation (8.3.3) can then be rewritten in the following form

$$D(k, p) = 1 + \frac{1}{(k\lambda_D)^2} [1 + \zeta Z(\zeta)] = 0. \quad (8.3.6)$$

If we define $\zeta = x + iy$, one can see from equation (8.3.2) and the definition $p = \gamma - i\omega$ that the dimensionless quantities x and y are proportional to the frequency and growth rate

$$x = \sqrt{\frac{m}{2\kappa T}} \frac{\omega}{k} \quad \text{and} \quad y = \sqrt{\frac{m}{2\kappa T}} \frac{\gamma}{k}. \quad (8.3.7)$$

To proceed further, it is necessary to develop appropriate power series expansions for $Z(\zeta)$. Two expansions are particularly useful. The first is the large-argument expansion ($|\zeta| \gg 1$), valid for high phase velocities, and the second is the small-argument expansion ($|\zeta| \ll 1$), valid for low phase velocities.

We start by discussing the large-argument expansion ($|\zeta| \gg 1$). For many applications the pole in the integration contour is located very close to the real z axis (i.e., $y \ll x$). This is normally the case for the large-argument expansion, since at high phase velocities the magnitude of the growth rate, $|\gamma|$, is usually very small. When the pole is located very close to the real z axis, the Plemelj relation can be used to evaluate the plasma dispersion function. Using the Plemelj relation, it is easy to show that the plasma dispersion function becomes

$$Z(\zeta) = i \frac{k}{|k|} \sqrt{\pi} e^{-\zeta^2} + \frac{1}{\sqrt{\pi}} P \int_{-\infty}^{\infty} \frac{e^{-z^2}}{z - \zeta} dz, \quad (8.3.8)$$

where the $k/|k|$ factor takes into account the sign change that occurs at $k = 0$. To evaluate the integral, one proceeds by expanding the $1/(z - \zeta)$ term as a series in inverse powers of ζ

$$\frac{1}{z - \zeta} = \frac{-1}{\zeta(1 - z/\zeta)} = -\frac{1}{\zeta} \left[1 + \left(\frac{z}{\zeta} \right) + \left(\frac{z}{\zeta} \right)^2 + \dots \right]. \quad (8.3.9)$$

Substituting this series into the integral, and integrating term by term, gives the following power series,

$$Z(\zeta) = i \frac{k}{|k|} \sqrt{\pi} e^{-\zeta^2} - \left[\frac{1}{\zeta} + \frac{1}{2\zeta^3} + \frac{3}{4\zeta^5} + \dots \right]. \quad (8.3.10)$$

For the small-argument expansion ($|\zeta| \ll 1$), an alternative representation must be used for $Z(\zeta)$, since the above expansion scheme does not converge if $\zeta/z \ll 1$. By simple substitution, it is straightforward to show that the plasma dispersion function satisfies the linear first-order differential equation,

$$\frac{1}{2} \frac{dZ}{d\zeta} + \zeta Z + 1 = 0. \quad (8.3.11)$$

This differential equation has several equivalent alternative solutions, one of which is easily shown to be

$$Z(\zeta) = i \frac{k}{|k|} \sqrt{\pi} e^{-\zeta^2} + i 2 e^{-\zeta^2} \int_0^{i\zeta} e^{-z^2} dz. \quad (8.3.12)$$

By expanding e^{-z^2} and $e^{-\zeta^2}$ in a power series, integrating term by term, and collecting the

results together in a single power series, it is easy to show that, for $|\zeta| \ll 1$, the plasma dispersion function can be represented by the expansion

$$Z(\zeta) = i \frac{k}{|k|} \sqrt{\pi} e^{-\zeta^2} - 2\zeta + \frac{4}{3}\zeta^3 - \frac{8}{15}\zeta^5 + \dots \quad (8.3.13)$$

By using equation (8.3.11) to eliminate the term $\zeta Z(\zeta)$ from equation (8.3.6), the dispersion relation can be written in the following convenient form

$$D(k, p) = 1 - \frac{1}{(k\lambda_D)^2} \frac{1}{2} Z'(\zeta) = 0, \quad (8.3.14)$$

where $Z'(\zeta)$ is the derivative of the plasma dispersion function. The derivative of the plasma dispersion function is easily obtained by simply differentiating the large-and small-argument expansions term by term, which for $|\zeta| \gg 1$ gives

$$Z'(\zeta) = -i \frac{k}{|k|} \sqrt{\pi} 2\zeta e^{-\zeta^2} + \frac{1}{\zeta^2} + \frac{3}{2} \frac{1}{\zeta^4} + \frac{15}{4} \frac{1}{\zeta^6} + \dots \quad (8.3.15)$$

and for $|\zeta| \ll 1$ gives

$$Z'(\zeta) = -i \frac{k}{|k|} \sqrt{\pi} 2\zeta e^{-\zeta^2} - 2 + 4\zeta^2 - \frac{8}{3}\zeta^4 + \dots \quad (8.3.16)$$

To illustrate the use of these expansions, we next evaluate the dispersion relation in the high-phase velocity limit. Using the first three terms of the large-argument expansion (8.3.15), substituting $\zeta = x + iy$, and expanding in powers of the small quantity y/x , the dispersion relation becomes

$$D(k, p) = 1 - \frac{1}{(k\lambda_D)^2} \left[\frac{1}{2x^2} + \frac{3}{4x^4} - i \left(\frac{y}{x^3} + \frac{k}{|k|} \sqrt{\pi} x e^{-x^2} \right) \right] = 0. \quad (8.3.17)$$

Setting the real and imaginary parts of the above equation equal to zero and using the definitions (8.3.7) for x and y , it is easy to show that the frequency and growth rate are given by

$$\omega^2 = \omega_p^2 + 3 \left(\frac{\kappa T}{m} \right) k^2, \quad (8.3.18)$$

and

$$\gamma = - \sqrt{\frac{\pi}{8}} \frac{\omega_p}{|k\lambda_D|^3} \exp \left[-\frac{1}{2(k\lambda_D)^2} - \frac{3}{2} \right], \quad (8.3.19)$$

where, as in Section 8.1.1, the successive approximation method has been used to solve for ω^2 . These results are identical to those obtained from the weak growth rate approximation, see equation (8.2.47).

8.4 The dispersion relation for a multi-component plasma

From inspection of the relevant equations, it is easy to see that the analysis of electrostatic waves in a hot multi-component plasma proceeds in a manner that is exactly the same as for a plasma of hot electrons and immobile ions, the only difference being that the dispersion relation now involves a sum over all species

$$D(k, p) = 1 - \sum_s \frac{\omega_{ps}^2}{k^2} \int_C \frac{\partial F_{s0}/\partial v_z}{v_z - ip/k} dv_z, \quad (8.4.1)$$

where the integration contour C must pass below the pole at $v_z = ip/k$ when k is positive, and above the pole when k is negative.

To illustrate the effects introduced by ions, it is convenient to limit our initial discussion to the case of electrons and a single species of positively charged ions. In this case, the sum in the dispersion can be simplified by taking advantage of the fact that $n_e = n_i$, which allows us to write

$$\sum_s \omega_{ps}^2 F_{s0}(v_z) = \omega_{pe}^2 \left[F_{e0}(v_z) + \frac{m_e}{m_i} F_{i0}(v_z) \right]. \quad (8.4.2)$$

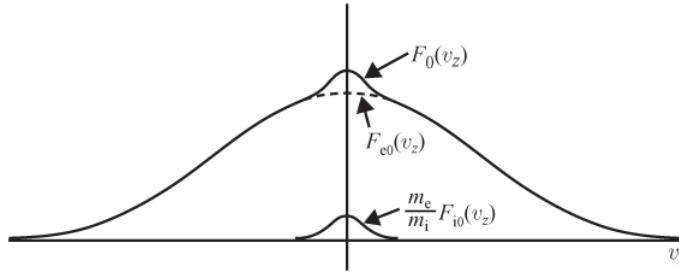


Fig. 8.16 The equivalent reduced distribution function, $F_0(v_z)$, for a two-component electron–ion plasma.

One can then define an equivalent reduced distribution function

$$F_0(v_z) = F_{e0}(v_z) + \frac{m_e}{m_i} F_{i0}(v_z). \quad (8.4.3)$$

Using the equivalent reduced distribution function, the dispersion relation can be written in a form that is identical to the dispersion relation for electrons and immobile ions

$$D(k, p) = 1 - \frac{\omega_{pe}^2}{k^2} \int_C \frac{\partial F_0/\partial v_z}{v_z - ip/k} dv_z. \quad (8.4.4)$$

It is interesting to note that because of the m_e/m_i factor, the contribution of the ions to the

equivalent distribution function is strongly suppressed. However, this does not mean that the ions are unimportant. A typical plot of the equivalent distribution function (for $T_e \approx T_i$) is shown in Figure 8.16. Note that the ions only produce a small bump in $F_0(v_z)$ near zero velocity. This bump becomes more sharply peaked as the ion temperature is decreased. As we shall soon see, the ratio of the electron temperature to the ion temperature is an important parameter in a hot electron-ion plasma.

8.4.1 Ion acoustic waves

Using the moment equations and the adiabatic equation of state, we demonstrated in Chapter ?? that ion motions lead to a new type of electrostatic wave called an ion acoustic wave. We next proceed to look for the roots of the multi-component dispersion relation that correspond to the ion acoustic wave.

8.4.1.1 Cauchy electron and ion velocity distributions

To simplify the analysis, we will first use Cauchy distributions for the zero-order electron and ion distribution functions

$$F_{e0}(v_z) = \frac{C_e}{\pi} \frac{1}{C_e^2 + v_z^2} \quad \text{and} \quad F_{i0}(v_z) = \frac{C_i}{\pi} \frac{1}{C_i^2 + v_z^2}. \quad (8.4.5)$$

Using the same procedure discussed in Section 8.2.3, the integrals in the dispersion relation can be evaluated using the residue theorem. It is then easy to show that the dispersion relation is given by

$$D(k, p) = 1 + \frac{\omega_{pe}^2}{(p + C_e|k|)^2} + \frac{\omega_{pi}^2}{(p + C_i|k|)^2} = 0. \quad (8.4.6)$$

To find the roots that correspond to the Langmuir mode we assume, as we did in Chapter ??, that the electron term is much larger than the ion term

$$\frac{\omega_{pe}^2}{(p + C_e|k|)^2} \gg \frac{\omega_{pi}^2}{(p + C_i|k|)^2}. \quad (8.4.7)$$

The dispersion relation then becomes

$$1 + \frac{\omega_{pe}^2}{(p + C_e|k|)^2} = 0, \quad (8.4.8)$$

which after substituting $p = \gamma - i\omega$ has solutions given by $\omega = \pm\omega_{pe}$ and $\gamma = -C_e|k|$, in agreement with our earlier results for Langmuir waves. One can easily demonstrate that the approximations used to obtain this solution are valid for any value of k . Thus, to a very good approximation, Langmuir waves are not affected by the presence of ions.

To find the roots that correspond to the ion acoustic mode, we assume that the electron

and ion terms in the dispersion relation are both much greater than one, so that we can omit the first term (i.e., the 1) in equation (8.4.6), which gives

$$\frac{\omega_{pe}^2}{(p + C_e|k|)^2} + \frac{\omega_{pi}^2}{(p + C_i|k|)^2} = 0. \quad (8.4.9)$$

After noting that $n_e = n_i$, the above equation simplifies to

$$(p + C_i|k|)^2 = -\frac{m_e}{m_i}(p + C_e|k|)^2. \quad (8.4.10)$$

Since we are looking for roots that have phase velocities much less than the electron thermal velocity, we assume that $p/|k| \ll C_e$ so that

$$p + C_i|k| = \pm i \left(\frac{m_e}{m_i} \right)^{1/2} C_e|k|, \quad (8.4.11)$$

or

$$\gamma - i\omega = \pm i \left(\frac{m_e}{m_i} \right)^{1/2} C_e|k| - C_i|k|. \quad (8.4.12)$$

If we use the minus sign, which gives the proper direction for the phase velocity, the phase velocity is given by

$$\frac{\omega}{k} = \pm \left(\frac{m_e}{m_i} \right)^{1/2} C_e, \quad (8.4.13)$$

where we have inserted \pm for $k/|k|$, and the growth rate is given by

$$\gamma = -C_i|k|. \quad (8.4.14)$$

It is readily verified that the assumptions used to obtain the above solutions are satisfied whenever $k\lambda_D \ll 1$. If we define the “temperature” of the Cauchy distributions by the relation $C_s^2 = \kappa T_s/m_s$, the phase velocity and growth rate for the ion acoustic wave simplify to

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e}{m_i}} \quad (8.4.15)$$

and

$$\gamma = -\sqrt{\frac{\kappa T_i}{m_i}}|k|. \quad (8.4.16)$$

As can be seen from equation (8.4.15), the phase velocity of the ion acoustic wave is proportional to the square root of the ratio of the electron temperature divided by the ion mass. This basic dependence is in agreement with our earlier results using the moment equations, see equation (??). However, equation (8.4.16) shows that the growth rate γ is

negative, which means the waves are damped. The ratio of the damping rate ($-\gamma$) to the wave frequency is given by

$$\frac{(-\gamma)}{\omega} = -\sqrt{\frac{T_i}{T_e}}. \quad (8.4.17)$$

As the above equation shows, if the electron and ion temperatures are comparable, the ion acoustic wave is strongly damped. This result is a marked departure from the results obtained using the moment equations, which indicated no damping at all. The damping of the ion acoustic wave is only small when $T_e \gg T_i$.

The strong dependence of the damping rate ($-\gamma$) on T_e/T_i has a simple interpretation. As shown by equation (8.4.15), the phase velocity of the ion acoustic wave is primarily dependent on the electron temperature. If $T_e = T_i$, the phase velocity is the same as the ion thermal velocity C_i , which means that the slope of the ion distribution function at the phase velocity of the wave is very large, thereby causing very strong Landau damping. The location of the phase velocity relative to the ion and electron distribution is shown in Figure 8.17. As the electron temperature is increased, the phase velocity increases and the damping decreases, since the slope of the ion distribution decreases. Except for extremely high phase velocities ($T_e \gg T_i$), electrons contribute very little to the damping. This is because the slope of the electron velocity distribution is very small at low velocities. In Figure 8.17 the ion thermal speed has been greatly exaggerated in relation to the electron thermal speed. For realistic electron-to-ion mass ratios and temperatures, the electron thermal speed is usually a factor of fifty to one hundred times larger than the ion thermal speed.

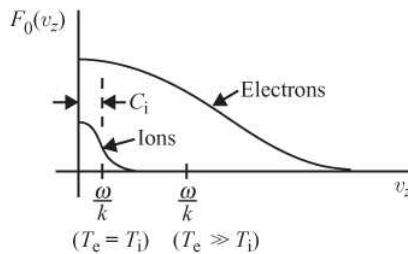


Fig. 8.17 When $T_e \gg T_i$ the ion acoustic wave is weakly damped because the phase velocity, ω/k , is much greater than the ion thermal speed, C_i . When $T_e = T_i$ the damping is very large because $\omega/k \approx C_i$.

8.4.1.2 Maxwellian electron and ion distributions

To gain experience using the plasma dispersion function, it is instructive to repeat the above analysis using Maxwellian distributions for the electrons and ions. By a simple extension of equation (8.3.14), it is easy to show that the dispersion relation for a multi-component Maxwellian plasma is given by

$$D(k, p) = 1 - \sum_s \frac{1}{(k\lambda_{Ds})^2} \frac{1}{2} Z'(\zeta_s) = 0, \quad (8.4.18)$$

where $Z'(\zeta_s)$ is the derivative of the plasma dispersion function, and ζ_s and λ_{Ds} are defined by

$$\zeta_s = \sqrt{\frac{m_s}{2\kappa T_s}} \left(\frac{ip}{k} \right) \quad \text{and} \quad \lambda_{Ds}^2 = \frac{\epsilon_0 \kappa T_s}{n_s e^2}. \quad (8.4.19)$$

For a plasma consisting of electrons and one species of positive ion, the dispersion relation can be rewritten in the following form

$$D(k, p) = 1 - \frac{1}{(k\lambda_{De})^2} \frac{1}{2} \left[Z'(\zeta_e) + \frac{T_e}{T_i} Z'(\zeta_i) \right] = 0, \quad (8.4.20)$$

where we have made use of the fact that $n_e = n_i$ and $\lambda_{De}^2 = (T_e/T_i)\lambda_{Di}^2$.

To obtain a completely general solution, the dispersion relation must be evaluated numerically. However, an approximate solution can be obtained for the ion acoustic mode if one assumes that the phase velocity is large compared to the ion thermal velocity and small compared to the electron thermal velocity. The large-argument expansion is then used for the ions

$$Z'(\zeta_i) = -i \frac{k}{|k|} \sqrt{\pi} 2 \zeta_i e^{-\zeta_i^2} + \frac{1}{\zeta_i^2}, \quad (8.4.21)$$

where we have only included the $1/\zeta_i^2$ term, and the small-argument expansion is used for the electrons

$$Z'(\zeta_e) = -i \frac{k}{|k|} \sqrt{\pi} 2 \zeta_e e^{-\zeta_e^2} - 2, \quad (8.4.22)$$

where we have only included the -2 term. Substituting these expansions into the dispersion relation, expanding $\zeta_i = x_i + iy_i$ in powers of y_i/x_i , and assuming $y_i \ll x_i$, the dispersion relation becomes

$$D(k, p) = 1 + \frac{1}{(k\lambda_{De})^2} \left\{ 1 - \frac{1}{2} \left(\frac{T_e}{T_i} \right) \frac{1}{x_i^2} \left(1 - i \frac{2y_i}{x_i} \right) + i \frac{k}{|k|} \sqrt{\pi} x_i \left[\sqrt{\frac{m_e}{m_i}} \sqrt{\frac{T_e}{T_i}} + \left(\frac{T_e}{T_i} \right) e^{-x_i^2} \right] \right\} = 0, \quad (8.4.23)$$

where we have made use of the relation $x_e = x_i \sqrt{m_e/m_i} \sqrt{T_i/T_e}$. Separating the real and imaginary parts, and writing x_i and y_i in terms of ω and γ then gives the following equations for the phase velocity and the growth rate,

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e}{m_i}} \frac{1}{(1 + k^2 \lambda_{De}^2)^{1/2}}, \quad (8.4.24)$$

and

$$\frac{\gamma}{\omega} = -\sqrt{\frac{\pi}{8}} \left[\sqrt{\frac{m_e}{m_i}} + \left(\frac{T_e}{T_i} \right)^{3/2} \exp \left(-\frac{T_e}{2T_i} \frac{1}{(1+k^2\lambda_{De}^2)^{3/2}} \right) \right] \times \frac{1}{(1+k^2\lambda_{De}^2)^{3/2}}. \quad (8.4.25)$$

Note that in the short wavelength limit, $k\lambda_{De} \gg 1$, the frequency has an upper limit given by the ion plasma frequency, $\omega_{pi} = (\sqrt{\kappa T_e/m_i})/\lambda_{De}$, in agreement with the results obtained in Section ?? from the moment equations (see Figure ??). In the long wavelength limit, $k\lambda_{De} \ll 1$, the equations for the phase velocity and growth rate simplify considerably and are given by

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e}{m_i}}, \quad (8.4.26)$$

and

$$\frac{\gamma}{\omega} = -\sqrt{\frac{\pi}{8}} \left[\sqrt{\frac{m_e}{m_i}} + \left(\frac{T_e}{T_i} \right)^{3/2} \exp \left(-\frac{T_e}{2T_i} \right) \right]. \quad (8.4.27)$$

Just as for the Cauchy distribution, the phase velocity is proportional to the square root of the electron temperature divided by the ion mass. Also, the damping rate ($-\gamma$) is very large when T_e is comparable to T_i . However, because of the exponential term, the damping decreases much more rapidly with increasing T_e/T_i than for a Cauchy distribution. The strong dependence on T_e/T_i is caused by the rapid decrease in the slope of the Maxwellian velocity distribution as the phase velocity increases. However, note that once the exponential term drops below $\sqrt{m_e/m_i}$, the damping no longer decreases as T_e/T_i increases. The $\sqrt{m_e/m_i}$ term is traceable to electron Landau damping via the -2 term in equation (8.4.22) and remains constant, because the phase velocity maintains a fixed relationship to the electron thermal speed.

Although the analytic results described above give a good approximation at large T_e/T_i , they do not give good results when the electron temperature is comparable to the ion temperature. Figure 8.18 shows a comparison of the damping rate ($-\gamma$)/ ω given by equation (8.4.27), labeled “to $1/\zeta_i^2$,” with a full numerical solution of the plasma dispersion. As can be seen, the agreement with the numerical solution is not very good for T_e/T_i values below about 20. Improved results can be obtained by keeping terms through $1/\zeta_i^4$ in the expansion for the ion plasma dispersion function. The phase velocity and growth rate are then given by

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e + 3\kappa T_i}{m_i}} \quad (8.4.28)$$

and

$$\frac{\gamma}{\omega} = -\sqrt{\frac{\pi}{8}} \frac{(3 + T_e/T_i)^{5/2}}{(9 + T_e/T_i)} \exp \left[-\frac{T_e}{2T_i} - \frac{3}{2} \right], \quad (8.4.29)$$

where again we have assumed that $k\lambda_{De} \ll 1$. Note that the phase velocity now has an additional term that is proportional to the ion temperature. This term accounts for the $\exp[-3/2]$ factor in the equation for the growth rate (8.4.29) and gives much better agreement with the exact numerical solution, as can be seen from the curve labeled “to $1/\lambda_i^4$ ” in Figure 8.18

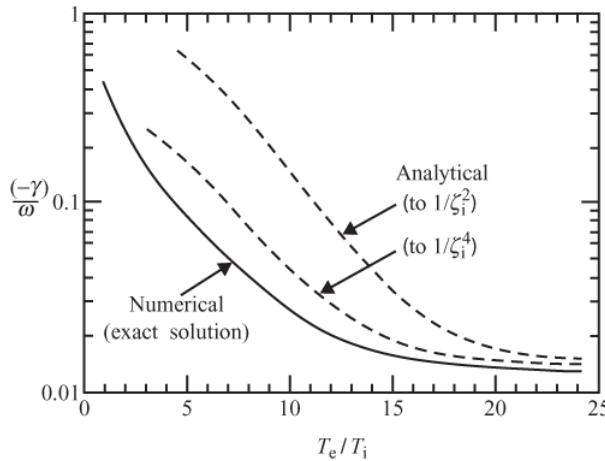


Fig. 8.18 A comparison of the approximate analytical expression for the damping, $(-\gamma)/\omega$, of ion acoustic waves with the results of an exact numerical calculation.

8.5 Stability

In the previous sections we have encountered a variety of both stable and unstable distribution functions. It is now of interest to consider the general conditions for stability. Since the growth or decay of a disturbance in a plasma is determined by the roots of the dispersion relation, it follows that the plasma is stable if all the roots have negative values for $\text{Re}\{p\}$, and unstable if one or more roots have positive values for $\text{Re}\{p\}$. The condition $\gamma = 0$ (i.e., $\text{Re}\{p\} = 0$) marks the boundary between stability and instability and is called the marginal stability condition.

8.5.1 Gardner's theorem

Before developing a general test for stability we first prove the following theorem due to Gardner [1963].

Theorem A single-humped velocity distribution is stable.

To prove this theorem, we start by assuming that the theorem is not true. It follows then that there are solutions to the dispersion relation for which $\text{Re}\{p\} > 0$. Since $\text{Re}\{p\} > 0$,

the integral in $D(k, p)$, see equation (8.2.21), can be expressed as an integral along the real velocity axis which, after substituting $p = \gamma - i\omega$, can be written

$$D(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{v_z - \omega/k - i\gamma/k} dv_z = 0. \quad (8.5.1)$$

Separating the integral into real and imaginary parts gives two equations,

$$D_r(k, p) = 1 - \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{(\partial F_0 / \partial v_z)(v_z - \omega/k)}{(v_z - \omega/k)^2 + (\gamma/k)^2} dv_z = 0, \quad (8.5.2)$$

and

$$D_i(k, p) = -\frac{\omega_p^2}{k^2} \frac{\gamma}{k} \int_{-\infty}^{\infty} \frac{\partial F_0 / \partial v_z}{(v_z - \omega/k)^2 + (\gamma/k)^2} dv_z = 0. \quad (8.5.3)$$

Next, suppose that $F_0(v_z)$ has a single hump with a maximum at velocity V_0 , as shown in Figure 8.19. Since D_r and D_i are both zero, any linear combination of these two terms is also zero. In particular, let us consider the combination

$$D_r - \left(\frac{kV_0 - \omega}{\gamma} \right) D_i = 0. \quad (8.5.4)$$

It is easy to show, after substituting for D_r and D_i , that the above equation simplifies to

$$1 + \frac{\omega_p^2}{k^2} \int_{-\infty}^{\infty} \frac{(\partial F_0 / \partial v_z)(V_0 - v_z)}{(v_z - \omega/k)^2 + (\gamma/k)^2} dv_z = 0. \quad (8.5.5)$$

Since F_0 has only a single hump, $\partial F_0 / \partial v_z$ is positive for $v_z < V_0$ and negative for $v_z > V_0$. Therefore, the integral is positive definite, which means that the above equation cannot be satisfied for any combination of ω and γ . Thus, we have arrived at a contradiction. The contradiction shows that the original assumption, $\text{Re}\{p\} > 0$, is false. Therefore, a single-humped distribution is always stable.

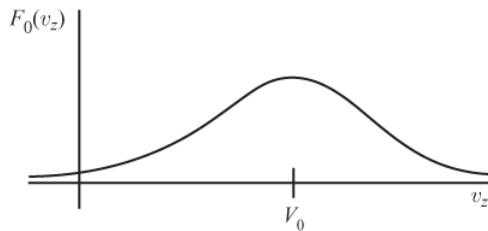


Fig. 8.19 | A velocity distribution with a single hump at velocity V_0 .

8.5.2 The Nyquist criterion

Having established that a single-humped velocity distribution is always stable, we next develop a general method for testing the stability of a distribution function with an arbitrary number of humps. This test is called the Nyquist criterion, after Nyquist [1932] who developed the method for analyzing the stability of electrical circuits. As shown earlier, the dispersion relation is a function of the complex frequency, $p = \gamma - i\omega$. Both γ and ω are functions of the wave number, k . The distribution function is unstable if some k exists for which $\gamma > 0$ and is stable if $\gamma < 0$ for all values of k . In the Nyquist approach, the procedure is to carry out a one-to-one mapping of the right half of the complex p -plane, $\text{Re}\{p\} = \gamma > 0$, onto the complex D -plane. If the mapping includes the point $D = 0$, then the plasma is unstable, since a solution of $D = 0$ exists for which $\gamma > 0$. If the mapping does not include the point $D = 0$ for any value of k , then the distribution function is stable.

To carry out the required mapping, it is useful to note that the boundary of the right half of the complex p -plane is the marginal stability line, $\gamma = 0$. The corresponding marginal stability line in the complex D -plane can then be determined by computing $D(k, p)$ along the $\gamma = 0$ line in the complex p -plane for $-\infty < \omega < +\infty$. The $\gamma = 0$ curve in the complex D -plane is always a closed curve. By mapping one additional point that is not on the $\gamma = 0$ line, it is a simple matter to determine whether the right half of the complex p -plane maps into the interior or the exterior of the $\gamma = 0$ curve.

To illustrate the Nyquist mapping procedure, consider a Cauchy velocity distribution function. From the dispersion relation for the Cauchy distribution, equation (8.2.31), it is easy to show that the real and imaginary parts of the dispersion relation, evaluated along the line $\gamma = 0$, are given by

$$D_r = 1 + \frac{\omega_p^2(C^2k^2 - \omega^2)}{(C^2k^2 + \omega^2)^2}, \quad (8.5.6)$$

and

$$D_i = 2 \frac{\omega_p^2 C |k| \omega}{(C^2k^2 + \omega^2)^2}. \quad (8.5.7)$$

As ω varies from minus infinity to plus infinity, these equations generate a closed cardioid-shaped curve in the complex D -plane, as shown in the right panel of Figure 8.20. It is easy to show that the right half of the complex p -plane (indicated by the shaded region in the left panel) maps into the interior of the cardioid-shaped curve. Note that the interior of the cardioid-shaped curve does not include the point $D = 0$. Since there is no solution of $D = 0$ for which γ is positive, it follows that the distribution function is stable. This is the expected result. Since a Cauchy velocity distribution has only a single hump, by Gardner's theorem the distribution function is stable.

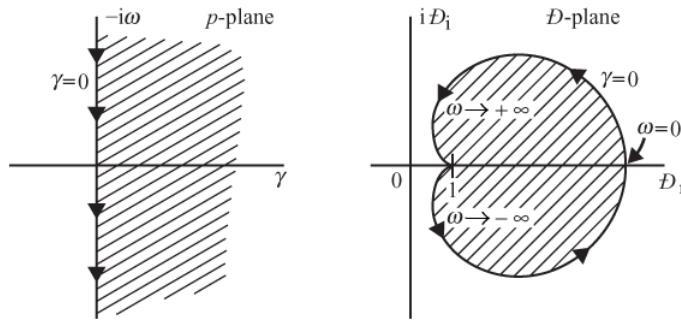


Fig. 8.20 A mapping of the right half of the complex p -plane onto the complex D -plane for a Cauchy distribution function.

8.5.3 The Penrose condition

Next we apply the Nyquist criterion to a distribution function with an arbitrary number of humps. To carry out the required mapping of the $\gamma = 0$ line onto the complex D -plane, we need an expression for $D(k, p)$ that is valid for a distribution function of arbitrary shape. Since γ is always small in the vicinity of the $\gamma = 0$ boundary, the Plemelj relation can be used to evaluate the real and imaginary parts of the dispersion relation, just as was done in the weak growth rate approximation (see Section 8.2.4). Using the Plemelj relation, it is easy to show that the real and imaginary parts of $D(k, p)$, evaluated along the $\gamma = 0$ line, are given by

$$D_r = 1 - \frac{\omega_p^2}{k^2} P \int_{-\infty}^{\infty} \frac{F'_0(v_z)}{v_z - \omega/k} dv_z, \quad (8.5.8)$$

and

$$D_i = -\pi \frac{k}{|k|} \frac{\omega_p^2}{k^2} F'_0\left(\frac{\omega}{k}\right). \quad (8.5.9)$$

To understand the geometry of the $\gamma = 0$ curve in the complex D -plane, we start by investigating the shape near $\omega = \pm\infty$. Throughout this analysis, we assume that the wave number k is positive. A similar analysis is easily repeated for negative k .

Since the slope of the distribution function, $F'_0(\omega/k)$, must go to zero as the velocity goes to infinity, it follows that $D = 1$ at $\omega = \pm\infty$. Furthermore, since the distribution function is always positive, $F'_0(\omega/k)$ must be negative as $\omega/k \rightarrow +\infty$, and positive as $\omega/k \rightarrow -\infty$. Therefore, the $\gamma = 0$ curve must approach $D = 1$ from above the D_r axis as $\omega/k \rightarrow +\infty$, and from below the D_r axis as $\omega/k \rightarrow -\infty$, as shown in Figure 8.21. Having established the shape of the $\gamma = 0$ curve near $D = 1$, we next explore the shape at intermediate frequencies. Since D_i can be zero only when $F'_0(\omega/k)$ is zero, it follows that the $\gamma = 0$ curve can only cross the D_r axis at those frequencies ω_j for which the derivative of the distribution function is zero (i.e., $\omega_j = V_j k$ where $F'(V_j) = 0$). For a smooth continuous

distribution function, there is always an odd number of such points. For a single-humped distribution function only one such point exists (for example, at $\omega = 0$ in Figure 8.20).

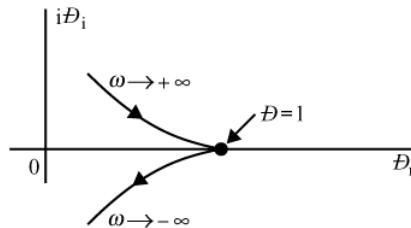


Fig. 8.21 An expanded scale representation of the mapping of the line $\gamma = 0$ onto the complex D -plane in the vicinity of $D = 1$.

To proceed further, we use a result from conformal mapping theory called the winding theorem [Flanigan, 1972]. The winding theorem states that if a closed contour C_p in the complex p -plane encloses n simple zeros of some mapping function $D(p)$, then the corresponding contour C_D in the complex D -plane must make n turns around the origin. To prove the winding theorem, we start by noting from the residue theorem that the number of turns a contour C_D makes around the origin ($D = 0$) is given by the integral

$$N_w = \frac{1}{2\pi i} \int_{C_D} \frac{dD}{D}. \quad (8.5.10)$$

The number N_w is called the winding number. According to the usual convention for complex integration, the winding number is positive for a contour that loops around the origin in the counter-clockwise sense. By a simple change of variables, the above integral can be converted to a corresponding integral in the complex p -plane

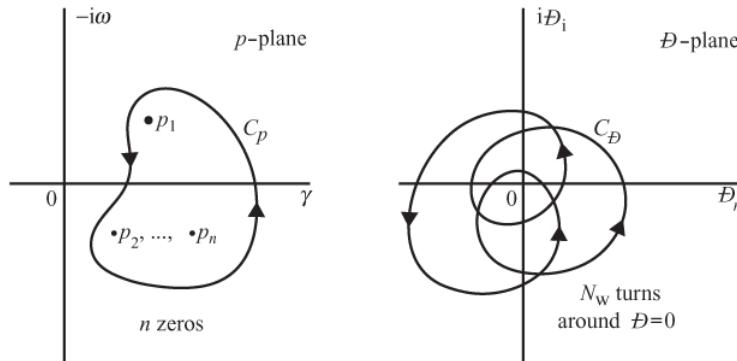


Fig. 8.22 If a contour encloses n roots, p_1, p_2, \dots, p_n of $D(p) = 0$ in the complex p -plane, then the mapping of this contour onto the complex D -plane must encircle the origin $N_w = n$ times.

$$N_w = \frac{1}{2\pi i} \int_{C_p} \frac{1}{D} \frac{\partial D}{\partial p} dp, \quad (8.5.11)$$

where the contour C_p is the corresponding mapping of the contour C_D onto the complex p -plane. The relationship between the two contours for a representative mapping function is shown in Figure 8.22. To understand the relationship between the number of zeros in the complex p -plane and the winding number, let us assume that the mapping function $D(p)$ has n zeros, p_1, p_2, \dots, p_n , within the contour C_p . Since the integral is independent of the shape of the contour, the contour can be deformed so that it shrinks closely around the n zeros. The integral can then be expressed as the sum of the integrals around each respective zero.

$$N_w = \sum_{j=1}^n \frac{1}{2\pi i} \int_{C_{pj}} \frac{1}{D} \frac{\partial D}{\partial p} dp. \quad (8.5.12)$$

Around each zero, the mapping function can be expanded in a Taylor series

$$D = 0 + \left. \frac{\partial D}{\partial p} \right|_{p_j} (p - p_j) + \dots, \quad (8.5.13)$$

where the first term is zero, since $D(p_j) = 0$. Similarly, the derivative of the mapping function can be expanded in a Taylor series, which to leading order is

$$\frac{\partial D}{\partial p} = \left. \frac{\partial D}{\partial p} \right|_{p_j} + \dots \quad (8.5.14)$$

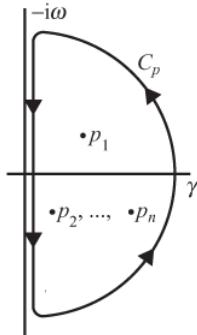


Fig. 8.23 The contour selected to count the total number of zeros of $D(p)$ in the right half of the complex p -plane.

Substituting (8.5.13) and (8.5.14) into equation (8.5.12), it is easy to show that

$$N_w = \sum_{j=1}^n \frac{1}{2\pi i} \int_{C_{pj}} \frac{dp}{p - p_j}. \quad (8.5.15)$$

From the residue theorem, each integral contributes a factor of $2\pi i$, so the sum yields $N_w = n$. Thus, we have shown that winding number is equal to the number of zeros within the contour, which proves the theorem.

To see how the winding theorem is used to investigate the geometry of the $\gamma = 0$ curve in the complex D -plane, we start by selecting a contour C_D that passes downward along the $\gamma = 0$ line in the complex p -plane. The contour is closed at infinity along the outer boundary of the right half-plane, as shown in Figure 8.23. The reason for this choice of the contour is that the integral $\int dD/D$ around the contour gives the total number of unstable roots. Note that any root of $D(k, p) = 0$ within this contour has $\text{Re}\{p\} > 0$, and is therefore unstable. Since the portion of the contour at infinity maps to the point $D = 1$, the rest of the contour maps onto the $\gamma = 0$ curve. From the winding theorem, we can now see that the number of times that the $\gamma = 0$ curve winds around the origin ($D = 0$) must be equal to the number of unstable roots of $D(k, p) = 0$. Since the number of unstable roots can never be negative, if the $\gamma = 0$ curve encircles the origin ($D = 0$), then it must wind around the origin at least once in the counter-clockwise sense.

We now have enough tools to determine the geometry of the $\gamma = 0$ curve in the complex D -plane. If the velocity distribution function has a single hump, then it can cross the D_r axis only once. Since the $\gamma = 0$ curve starts from below the D_r axis at $\omega = -\infty$, this crossing must be in the upward direction, toward increasing D_i . Since the winding number must be positive, the curve must cross the D_r axis to the right of $D = 0$, otherwise the curve would encircle the origin in the clockwise sense, which is prohibited. Since the crossing of the D_r axis must be to the right of $D = 0$, the curve does not encircle the origin, so the distribution function is stable. Thus, we have again proven Gardner's theorem, which states that a single-humped velocity distribution is always stable.

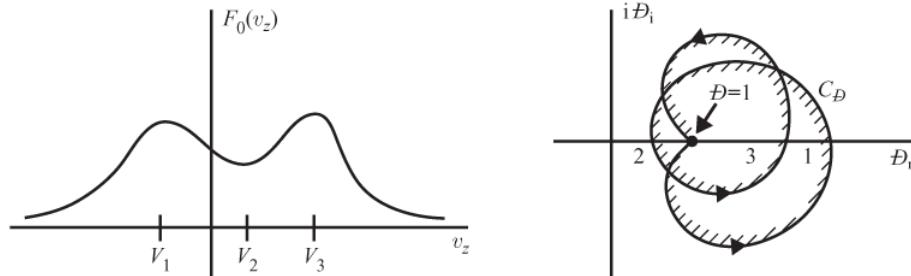


Fig. 8.24 The mapping of the line $\gamma = 0$ onto the complex D -plane for a double-humped velocity distribution.

If the velocity distribution function has two humps, the $\gamma = 0$ curve crosses the D_r axis three times, corresponding to the three velocities, V_1 , V_2 , and V_3 , for which $F'_0(V_j) = 0$. Two of these crossings, V_1 and V_3 , are in the upward direction, toward positive D_i , and correspond to local maxima in $F_0(v_z)$. The remaining crossing, V_2 , is in the downward direction and corresponds to a local minimum in $F_0(v_z)$. A representative $\gamma = 0$ curve, and its relationship to $F_0(v_z)$ is shown in Figure 8.24. From equation (8.5.8), it is easy to see that the D_r values at which the $\gamma = 0$ line crosses the D_r axis are given by

$$D_r = 1 - \frac{\omega_p^2}{k^2} P \int_{-\infty}^{\infty} \frac{F'_0(v_z)}{v_z - V_j} dv_z, \quad (8.5.16)$$

where the V_j are the velocities for which $F'_0(V_j) = 0$. Since the right half of the complex p -plane maps into the interior of the $\gamma = 0$ curve, if any one of these crossings occurs to the left of the origin (i.e., $D_r < 0$), then by the Nyquist criterion the distribution function is unstable. Since the wave number k can take on any value whatever, including very small values, it is easy to see that the condition for a negative D_r value at a crossing of the D_r axis is that the integral

$$P \int_{-\infty}^{\infty} \frac{F'_0(v_z)}{v_z - V_j} dv_z > 0, \quad (8.5.17)$$

be positive, where the V_j are the velocities at which $F'_0(V_j) = 0$. Therefore, the existence of a positive value for the above integral for at least one value of V_j is a necessary condition for instability. Since $F_0(V_j)$ is a constant, it is easy to see that the above integral can be rewritten in the form

$$P \int_{-\infty}^{\infty} \frac{F'_0(v_z)}{v_z - V_j} dv_z = P \int_{-\infty}^{\infty} \frac{\frac{\partial}{\partial v_z}[F_0(v_z) - F_0(V_j)]}{v_z - V_j} dv_z > 0. \quad (8.5.18)$$

The integral on the right-hand side can be integrated by parts once to give the following condition for instability

$$\int_{-\infty}^{\infty} \frac{F_0(v_z) - F_0(V_j)}{(v_z - V_j)^2} dv_z > 0. \quad (8.5.19)$$

The above relation is called the Penrose condition, after Penrose [1960] who first showed that it is a necessary and sufficient condition for instability. Note that since the numerator in the integrand goes to zero at $v_z = V_j$, it is not necessary to specify that the integral be a principal value. It is also easy to see that the Penrose condition also applies for an arbitrary number of humps in the velocity distribution function.

The Penrose condition has a simple graphical interpretation. Consider the double-humped velocity distribution function shown in Figure 8.25. The integral in the Penrose condition gives the area above (+) and below (-) the line $F_0 = F_0(V_j)$, weighted by the function $1/(v_z - V_j)^2$. From Figure 8.25 one can see that for instability to occur it is not sufficient that the distribution be simply double humped. The two humps must be large enough that the positive contributions to the integral exceed the negative contributions.

8.5.4 Some representative instabilities

To illustrate the principles discussed above, we next consider various types of potentially unstable distribution functions. In all cases the instabilities involve distribution functions that have a finite temperature. In this respect they differ fundamentally from the cold, delta function, beam instabilities discussed earlier in this chapter. Whereas the cold beam instabilities could equally well have been analyzed using the fluid equation, the introduction of a finite temperature leads to an entirely new class of instabilities, called kinetic instabilities, that depend on the detailed shape of the distribution function.

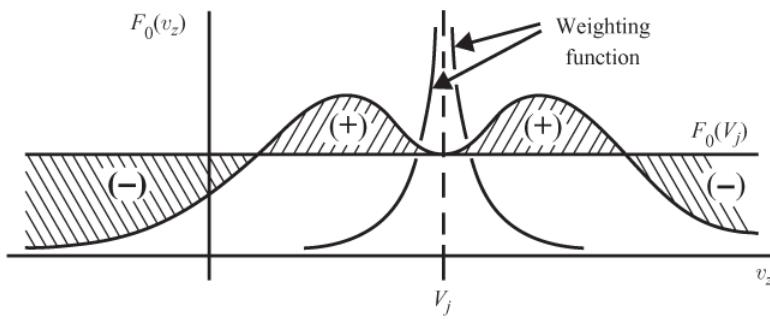


Fig. 8.25 A schematic interpretation of the Penrose criterion.

8.5.4.1 The counter-streaming beam instability

A simple example that illustrates many of the above principles is provided by the counter-streaming Cauchy distribution

$$F_0(v_z) = \frac{C}{2\pi} \left[\frac{1}{C^2 + (v_z - V)^2} + \frac{1}{C^2 + (v_z + V)^2} \right]. \quad (8.5.20)$$

In the limit of zero “temperature” (i.e., $C = 0$), this distribution reduces to two oppositely directed delta function beams, which from our earlier analysis is known to be unstable. As the temperature increases, the beams spread in velocity, eventually turning into a single-humped distribution, which from Gardner’s theorem is known to be stable. It is easy to show that the transition from a double-humped to a single-humped distribution takes place when $C = \sqrt{3V}$. Thus, the transition from unstable to stable must occur at some point between $C = 0$ and $C = \sqrt{3V}$.

The criterion for the onset of instability can be obtained by direct evaluation of the Penrose condition. There are three points where $F'_0(V_j) = 0$, near the two peaks associated with $v_z = \pm V$, and at $V_j = 0$. The points near the two peaks do not produce instability since the integral in the Penrose condition is in both cases clearly negative. For $V_j = 0$, the integral in the Penrose condition can be evaluated directly and is given by

$$\int_{-\infty}^{\infty} \frac{F_0(v_z) - F_0(V_j)}{(v_z - V_j)^2} dv_z = \frac{V^2 - C^2}{(V^2 + C^2)^2}. \quad (8.5.21)$$

The plasma is unstable whenever the above integral is positive, which occurs when $V > C$. The shape of the distribution function at the onset of instability ($V = C$) is shown by the solid curve in Figure 8.26.

By direct evaluation of D_r and D_i , it is easy to show that the Nyquist plot for $C < V$ is as shown in Figure 8.27. The $\gamma = 0$ curve loops around the origin once which, by the Nyquist theorem, indicates the presence of an unstable root in the dispersion relation.

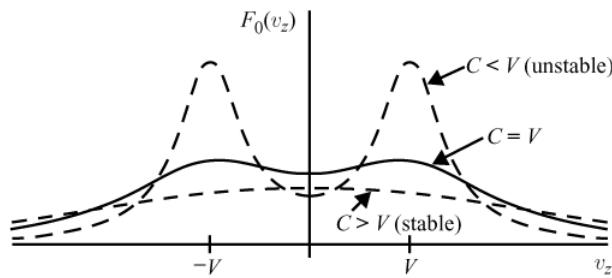


Fig. 8.26 For a pair of counter-streaming Cauchy distributions with thermal velocity C and drift velocities V and $-V$, the plasma is stable if $C > V$, marginally stable if $C = V$, and unstable if $C < V$.

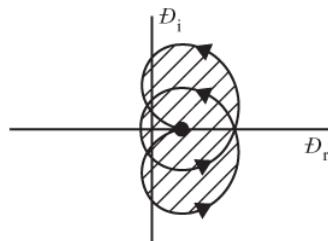


Fig. 8.27 The mapping of the right half of the complex p -plane onto the complex D -plane for a counter-streaming Cauchy distribution with $C < V$. Since the mapping includes the origin ($D = 0$), the plasma is unstable.

8.5.4.2 The “bump-on-tail” instability

Using the weak growth rate approximation, we previously showed that the growth rate γ at high phase velocities is proportional to the slope of the distribution function evaluated at the phase velocity of the wave. Therefore, any distribution function that has a region of positive slope at high phase velocities leads to exponentially growing waves. Since the weak growth rate approximation is only valid at high phase velocities, well above the thermal velocity ($\omega/k \gg v_{\text{th}}$), this analysis is applicable if the region of positive slope lies in the tail of the distribution function, well above the thermal velocity, such as shown in Figure 8.28. This type of instability is often called a “bump-on-tail” instability. Since the growth rate is proportional to $\partial F_0 / \partial v_z$, it is easy to see that a range of phase velocities slightly below the peak is unstable. If the bump-on-tail overlaps with the main thermal distribution, the growth rate is made up of two terms, one negative (damping), caused by the tail of the main thermal distribution, and the other positive (growth), due to the positive slope of the bump. Thus, the overall growth rate is a competition between two effects. If the velocity at which the bump occurs gets too low, near the main part of the thermal distribution, or if the positive slope on the bump gets too small, the instability ceases. Notice that for a bump-on-tail distribution the growth rate is directly proportional to the density of the bump, $\gamma \propto n_b$. This dependence differs from the cold weak beam approximation discussed earlier, which has a growth rate proportional to $n_b^{1/3}$. The reason for this difference is that

the physical mechanisms involved in the two instabilities are different. The bump-on-tail instability involves only those particles that are in resonance with the wave. Particles with velocities greater than the phase velocity contribute to the growth of the wave, and those with velocities less than the phase velocity contribute to damping. This type of behavior is typical of a kinetic instability. The cold weak beam instability, on the other hand, involves a fluid-like motion of the beam, with all of the particles oscillating in unison. This type of behavior is typical of a fluid instability.

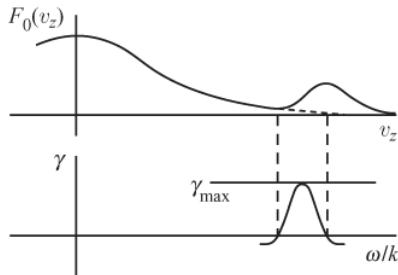
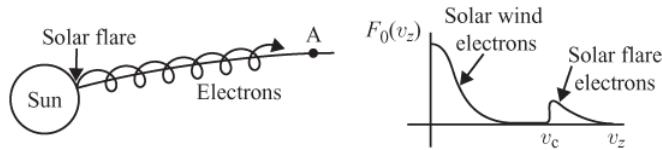
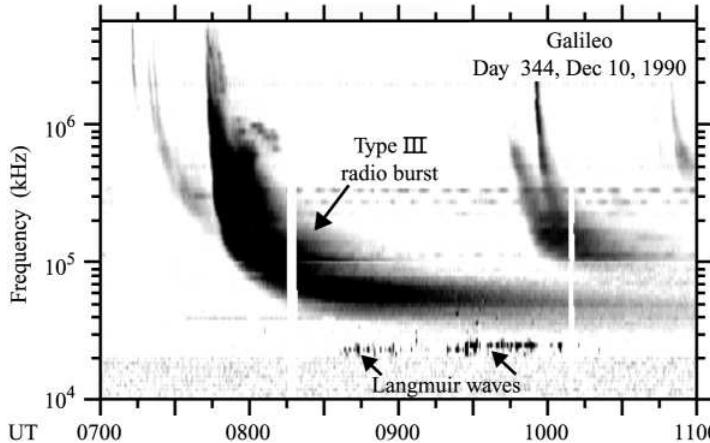


Fig. 8.28 A bump-on-tail distribution produces growing waves over a narrow range of phase velocities, ω/k , where $F_0(v_z)$ has a positive slope.

Numerous examples of electrostatic instabilities produced by a bump-on-tail distribution can be found in both laboratory and space plasmas. A good example in a space plasma is given by a type III solar radio burst, which is produced by electrons emitted from the Sun by a solar flare. When the flare occurs, a very broad range of electron energies (up to several hundred keV) are impulsively released from the Sun as shown in Figure 8.29. These electrons stream outward from the Sun along the solar wind magnetic field lines. At a given time after the flare, only those electrons with velocities greater than some cutoff velocity v_c can reach point A along the magnetic field line. Solar flare electrons of lower velocity cannot reach this point because the transit time is too long. This velocity cutoff automatically causes a bump-on-tail type of distribution function and leads to the growth of Langmuir waves (electron plasma oscillations) near the local electron plasma frequency. As the solar flare electron distribution evolves, a localized region of plasma oscillations is produced that moves outward from the Sun. Since the plasma oscillations are generated near the Sun, they cannot be detected directly from the Earth. However, the plasma oscillations produce electromagnetic radiation that can be detected at the Earth. The radiation is generated by non-linear interactions at the plasma frequency and at twice the plasma frequency [Ginzburg and Zhelezniakov, 1958]. Since the electron plasma frequency decreases with increasing radial distance from the Sun, the frequency of the radio emission decreases with increasing time as the region of plasma oscillations moves outward from the Sun. This frequency-time variation is a characteristic feature of type III radio bursts. A spectrogram of an intense type III solar radio burst and its associated Langmuir waves is shown in Figure 8.30.

**Fig. 8.29**

Electrons ejected by a solar flare produce a bump-on-tail distribution because of time of flight considerations, thereby causing the growth of Langmuir waves in the solar wind.

**Fig. 8.30**

An example of a type III solar radio burst and its associated Langmuir waves [from Gurnett et al., 1993].

8.5.4.3 The current-driven ion acoustic instability

When a current is present in a plasma, the electrons have a net drift relative to the ions. If the relative drift between the electrons and the ions is sufficiently large, a double hump occurs in the equivalent reduced distribution function, as shown in Figure 8.31. The ion acoustic mode can then be driven unstable. This instability is called the current-driven ion acoustic instability.

To provide a quantitative evaluation of the threshold current required to drive the ion acoustic instability, we first consider Cauchy distributions for the electrons and ions. To produce a current, the drift velocity of the electrons is assumed to be U_e , and the drift velocity of the ions is assumed to be zero. The appropriate distribution functions are then given by

$$F_{e0}(v_z) = \frac{C_e}{\pi} \frac{1}{C_e^2 + (v_z - U_e)^2}, \quad (8.5.22)$$

and

$$F_{i0}(v_z) = \frac{C_i}{\pi} \frac{1}{C_i^2 + v_z^2}. \quad (8.5.23)$$

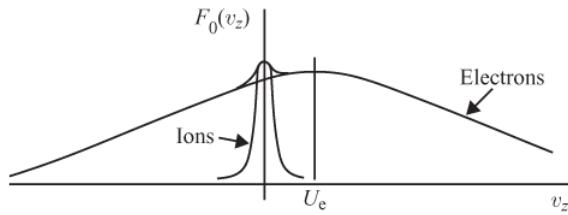


Fig. 8.31 When a current is present, a double hump is produced in the equivalent reduced velocity distribution function. This double hump can lead to the growth of ion acoustic waves.

Following the same procedure described in Section 8.2.3, it is easy to show that the dispersion relation is given by

$$D(k, p) = 1 + \frac{\omega_{pe}^2}{(p + ikU_e + C_e|k|)^2} + \frac{\omega_{pi}^2}{(p + C_i|k|)^2} = 0, \quad (8.5.24)$$

where the second term is due to the drifting electron distribution and the third term is due to the ion distribution. The kU_e term in the denominator of the electron term arises when the residue is evaluated at $v_z = U_e - iC_e$, and is simply the Doppler shift produced by the electron drift. As before, for $k\lambda_{De} \gg 1$, the ion acoustic root is obtained by assuming that the electron and ion terms are much larger than one so that

$$\frac{\omega_{pe}^2}{(p + ikU_e + C_e|k|)^2} = \frac{-\omega_{pi}^2}{(p + C_i|k|)^2}. \quad (8.5.25)$$

Solving the above equation for p then gives

$$p + C_i|k| = \pm i \left(\frac{m_e}{m_i} \right)^{1/2} (p + ikU_e + C_e|k|), \quad (8.5.26)$$

which can be rewritten in the form

$$p \left(1 \mp i \left(\frac{m_e}{m_i} \right)^{1/2} \right) = - \left(C_i \pm \left(\frac{m_e}{m_i} \right)^{1/2} U_e \frac{k}{|k|} \right) |k| \pm i \left(\frac{m_e}{m_i} \right)^{1/2} C_e |k|. \quad (8.5.27)$$

Using the minus sign, which assures that the phase velocity is in the proper direction, and neglecting the $i\sqrt{m_e/m_i}$ term on the left, which is small, it is easy to show that the phase velocity is given by

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e}{m_i}}, \quad (8.5.28)$$

where we used $C_e^2 = \kappa T_e/m_e$ and inserted \pm for $k/|k|$. The corresponding growth rate is given by

$$\gamma = - \left(C_i - U_e \frac{k}{|k|} \left(\frac{m_e}{m_i} \right)^{1/2} \right) |k|. \quad (8.5.29)$$

As can be seen, if the electron drift velocity, U_e , is sufficiently large, the growth rate can become positive, indicating instability. Note that the unstable waves occur for k in the same direction as the electron drift (k and U_e must have the same sign). The threshold drift velocity at which the instability starts is given by

$$|U_e^*| = \sqrt{\frac{m_i}{m_e}} C_i = \left(\frac{T_i}{T_e} \right)^{1/2} C_e. \quad (8.5.30)$$

This equation shows that if the electron and ion temperatures are the same ($T_e = T_i$), the threshold drift velocity is the electron thermal speed, which is usually very large. Therefore, a very large current is required to drive the ion acoustic mode unstable. The instability threshold can be reduced considerably if the electron temperature is increased relative to the ion temperature, i.e., $T_e \gg T_i$. This reduction is caused by the decreased ion Landau damping that occurs when $T_e \gg T_i$ (see Section sec8.4.1.1.1).

The ion acoustic instability can also be analyzed for a Maxwellian distribution of electron and ion velocities. Using a procedure similar to that used in Section sec8.4.1.2, it is straightforward, but somewhat tedious, to show that the growth rate in the long wavelength limit ($k\lambda_{De} \ll 1$) is given by

$$\frac{\gamma}{\omega} = - \sqrt{\frac{\pi}{8}} \left[\sqrt{\frac{m_e}{m_i}} \left(1 - \frac{k}{|k|} \frac{U_e}{C_A} \right) + \left(\frac{T_e}{T_i} \right)^{3/2} \exp\left(-\frac{T_e}{2T_i}\right) \right], \quad (8.5.31)$$

where $C_A = \sqrt{\kappa T_e / m_i}$ is the ion acoustic speed. In the derivation of this equation, the plasma dispersion function for the ion term has been expanded to order $1/\zeta_i^2$, as in equation (8.4.21). The threshold electron drift velocity, U_e^* , at which the instability starts, is obtained by setting the term in the large brackets to zero and solving for U_e , which gives

$$|U_e^*| = \sqrt{\frac{\kappa T_e}{m_e}} \left[\sqrt{\frac{m_e}{m_i}} + \left(\frac{T_e}{T_i} \right)^{3/2} \exp\left(-\frac{T_e}{2T_i}\right) \right]. \quad (8.5.32)$$

If $T_e = T_i$, the threshold is again very large, on the order of the electron thermal speed, but decreases rapidly as T_e/T_i increases. A comparison of the threshold drift velocity given by equation (8.5.32) with the results from a full numerical solution is shown as a function of T_e/T_i in Figure 8.32. As can be seen, equation (8.5.32) provides a good approximation for large T_e/T_i , but greatly over-estimates the threshold at low T_e/T_i values. The discrepancy between the analytic and numerical results at low T_e/T_i values can be reduced considerably by keeping terms to order $1/\zeta_i^4$ in the expansion for the ion plasma dispersion function, similar to the derivation of equations (8.4.28) and (8.4.29). This exercise is left to the reader.

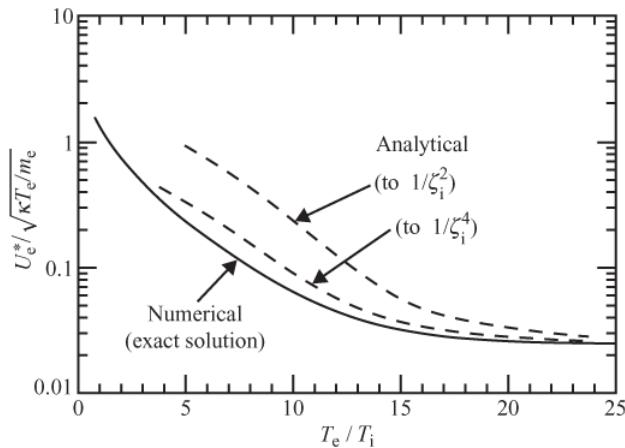


Fig. 8.32 A comparison of the approximate analytical solution for the threshold electron drift velocity required to produce unstable ion acoustic waves with the results of an exact numerical calculation.

Problems

8.1. A plasma has a “spherical shell” distribution function given by

$$f_0(\mathbf{v}) = \frac{n_0}{4\pi c^2} \delta(|\mathbf{v}| - c)$$

where c is a constant.

- (a) Using the Fourier analysis approach, show that the dispersion relation for electrostatic waves in this plasma is $\omega^2 = \omega_p^2 + c^2 k^2$.

(b) What is the region of validity of this dispersion relation?

8.2. A plasma consists of two homogeneous streams of electrons with equal densities and equal but opposite velocities ($V_1 = -V_2$).

- (a) Show that the critical wave number for electrostatic instability is

$$k_c = \sqrt{2} \frac{\omega_{p1}}{V_1}$$

- (b) Show that the maximum growth rate is

$$\gamma = \frac{1}{2} \omega_{p1}$$

8.3. Buneman [1959] analyzed the maximum growth rate of an electron beam streaming at velocity \mathbf{V} through an equal density of ions initially at rest. The dispersion relation for this system is

$$D(k, \omega) = 1 - \frac{\omega_{pi}^2}{\omega^2} - \frac{\omega_{pe}^2}{(kV - \omega)^2} = 0.$$

(a) If we assume $\omega_{\text{pi}}^2 \ll \omega^2$, show that the dispersion relation can be written

$$kV = \omega + \omega_{\text{pe}} + \frac{1}{2} \frac{\omega_{\text{pe}} \omega_{\text{pi}}^2}{\omega^2}.$$

(b) Show that the above dispersion relation has a solution for real k of the form

$$\omega = (\omega_{\text{pe}} \omega_{\text{pi}}^2 \cos \theta)^{1/3} e^{i\theta}$$

where θ is any arbitrary angle.

(c) Show that the imaginary part of ω reaches a maximum when $\theta = \pi/3$.

(d) Show that, when $\theta = \pi/3$,

$$\gamma = \omega_{\text{pe}} \left(\frac{m_e}{2m_i} \right)^{1/3} \frac{\sqrt{3}}{2} \quad \text{and} \quad \omega_r = \omega_{\text{pe}} \left(\frac{m_e}{2m_i} \right)^{1/3} \frac{1}{2}.$$

(e) Is assumption (a) satisfied?

8.4. The Laplace transform of $f(t)$ is defined by

$$\tilde{f}(p) = \int_0^\infty f(t) e^{-pt} dt.$$

Show that the Laplace transform of $f(t) = \cosh(at)$ is given by

$$\tilde{f}(p) = \frac{p}{p^2 - a^2}.$$

8.5. The inverse Laplace transform of $\tilde{f}(p)$ is given by

$$f(t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} \tilde{f}(p) e^{pt} dp,$$

where the integration along the line $p = \sigma$ must pass to the right of any poles of $\tilde{f}(p)$. Use the residue theorem to evaluate the inverse Laplace transform of

$$\tilde{f}(p) = \frac{1}{p^2 - a^2}.$$

8.6. A plasma has a reduced one-dimensional distribution given by the Cauchy distribution

$$F_0(v_z) = \frac{C}{\pi} \frac{1}{C^2 + v_z^2}.$$

(a) Using the Laplace transform approach, show that the dispersion relation is

$$D(k, p) = 1 + \frac{\omega_p^2}{(p + C|k|)^2} = 0.$$

- (b) Show that the roots of this dispersion relation (with $p = \gamma - i\omega$) give $\gamma = -C|k|$ and $\omega = \pm\omega_p$.

8.7. In the previous problem we showed that the Cauchy distribution

$$F_0(v_z) = \frac{C}{\pi} \frac{1}{v_z^2 + C^2},$$

gives a dispersion relation

$$D(k, p) = 1 + \frac{\omega_p^2}{(p + C|k|)^2},$$

which has a solution $\omega = \pm\omega_p$ and $\gamma = -C|k|$. Show that in the high phase velocity limit (i.e., $|k| \rightarrow 0$), the weak growth rate approximation,

$$D_r(k, \omega) = 0 \quad \text{and} \quad \gamma = \frac{-D_i}{\partial D_r / \partial \omega},$$

gives the same result.

8.8. To model a hot beam, one can use a shifted Cauchy distribution of the form

$$F_0(v_z) = \frac{C}{\pi} \frac{1}{C^2 + (v_z - U)^2}$$

where U is the beam velocity. Show that the dispersion relation for this plasma is

$$D(k, p) = 1 + \frac{\omega_p^2}{(p + C|k| + ikU)^2} = 0.$$

8.9. The reduced one-dimensional kappa distribution is given by

$$F_0(v_z) = \frac{\Gamma(\kappa + 1)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\kappa + \frac{1}{2}\right)\kappa^{1/2}C} \left[1 + \frac{v_z^2}{\kappa C^2}\right]^{-(\kappa+1)}$$

where Γ is the gamma function and C is the thermal velocity.

- (a) Show that in the limit $\kappa \rightarrow \infty$, the kappa distribution becomes a Gaussian distribution.

- (b) Show that for small κ , the kappa distribution approaches a Cauchy distribution.

8.10. The plasma dispersion function is defined by the equation

$$Z(\zeta) = \frac{1}{\sqrt{\pi}} \int_C \frac{e^{-z^2}}{z - \zeta} dz,$$

where the contour C runs along the real ζ axis, and passes around the pole at $\zeta = z$ when $\text{Re}\{\zeta\} < 0$. Explain why this integral cannot be evaluated by extending the contour around the upper half-plane in the same way that was done for the Cauchy distribution function.

- 8.11. If the pole at $\zeta = x + iy$ is very close to the z axis (i.e., $|y| \ll |x|$), the plasma dispersion function is given by

$$Z(\zeta) = i \frac{k}{|k|} \sqrt{\pi} e^{-\zeta^2} + \frac{1}{\sqrt{\pi}} P \int_{-\infty}^{\infty} \frac{e^{-z^2}}{z - \zeta} dz.$$

By writing

$$\frac{1}{z - \zeta} = \frac{-1}{\zeta(1 - z/\zeta)} = -\frac{1}{\zeta} \left[1 + \left(\frac{z}{\zeta} \right) + \left(\frac{z}{\zeta} \right)^2 + \dots \right]$$

and integrating term by term, show that in the limit of large ζ the plasma dispersion function is given by the following power series

$$Z(\zeta) = i \frac{k}{|k|} \sqrt{\pi} e^{-\zeta^2} - \left[\frac{1}{\zeta} + \frac{1}{2\zeta^3} + \frac{3}{4\zeta^5} + \dots \right].$$

- 8.12. It is easy to show that the plasma dispersion function, $Z(\zeta)$, is the solution of the following first-order differential equation

$$\frac{dZ}{d\zeta} + 2\zeta Z = -2.$$

Show that the following functions satisfy this differential equation

- (a) $Z(\zeta) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{e^{-z^2}}{z - \zeta} dz$
- (b) $Z(\zeta) = 2ie^{-\zeta^2} \int_{-\infty}^{i\zeta} e^{-z^2} dz$
- (c) $Z(\zeta) = i\sqrt{\pi}e^{-\zeta^2} [1 + \operatorname{erf}(i\zeta)]$
where

$$\operatorname{erf}(i\zeta) = \frac{2}{\sqrt{\pi}} \int_0^{i\zeta} e^{-z^2} dz.$$

- 8.13. Using the erf representation for $Z(\zeta)$ in (b) above, show that for small ζ ,

$$Z(\zeta) = i\sqrt{\pi}e^{-\zeta^2} - 2\zeta + \frac{4}{3}\zeta^3 - \frac{8}{15}\zeta^5 + \dots$$

Hint, use a Taylor series expansion for $\exp[-\zeta^2]$ and $\exp[-z^2]$, integrate term by term, and then collect like powers of ζ .

- 8.14. For a plasma consisting of electrons and one species of positive ion, both with Maxwellian velocity distributions, the dispersion relation can be written

$$D(k, p) = 1 - \frac{1}{(k\lambda_{De})^2} \frac{1}{2} \left[Z'(\zeta_e) + \frac{T_e}{T_i} Z'(\zeta_i) \right] = 0,$$

where T_e is the electron temperature and T_i is the ion temperature. Use the large-argument expansion of the plasma dispersion function for the ions

$$Z'(\zeta_i) = -i \frac{k}{|k|} \sqrt{\pi} 2 \zeta_i e^{-\zeta_i^2} + \frac{1}{\zeta_i^2},$$

and the small-argument expansion for the electrons

$$Z'(\zeta_e) = -i \frac{k}{|k|} \sqrt{\pi} 2 \zeta_e e^{-\zeta_e^2} - 2.$$

Substitute these expansions into the dispersion relation, expand in powers of y_i/x_i assuming $y_i \ll x_i$, and show that

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e}{m_i}} \frac{1}{(1 + k^2 \lambda_{De}^2)^{1/2}},$$

and

$$\frac{\gamma}{\omega} = - \sqrt{\frac{\pi}{8}} \left[\sqrt{\frac{m_e}{m_i}} + \left(\frac{T_e}{T_i} \right)^{3/2} \exp \left(-\frac{T_e}{2T_i} \frac{1}{(1 + k^2 \lambda_{De}^2)} \right) \right] \times \frac{1}{(1 + k^2 \lambda_{De}^2)^{3/2}}.$$

8.15. A Maxwellian plasma consists of equal densities of electrons and ions with temperatures T_e and T_i . The dispersion relation for this plasma can be written

$$D(k, p) = 1 - \frac{1}{(k \lambda_{De})^2} \frac{1}{2} Z'(\zeta_e) - \frac{1}{(k \lambda_{Di})^2} \frac{1}{2} Z'(\zeta_i) = 0.$$

By using the low-phase velocity expansion for the electrons ($\zeta_e \ll 1$) and the high-phase velocity expansion for the ions ($\zeta_i \gg 1$), and keeping all terms through $1/\zeta_i^4$ in the expansion for $Z(\zeta_i)$ and only the -2 term for $Z(\zeta_e)$, show that if $k \lambda_{De} \ll 1$ and $T_e \gg T_i$, then to a good approximation

$$\frac{\omega}{k} = \pm \sqrt{\frac{\kappa T_e + 3\kappa T_i}{m_i}}$$

and

$$\frac{\gamma}{\omega} = - \sqrt{\frac{\pi}{8}} \frac{(3 + T_e/T_i)^{5/2}}{(9 + T_e/T_i)} \exp \left[-\frac{T_e}{2T_i} - \frac{3}{2} \right].$$

8.16. A plasma consists of two electron beams of equal “temperature” and density moving with velocities U and $-U$. The reduced one-dimensional distribution function is given by a counter-streaming Cauchy distribution

$$F_0(v_z) = \frac{C}{2\pi} \left[\frac{1}{C^2 + (v_z - U)^2} + \frac{1}{C^2 + (v_z + U)^2} \right].$$

(a) Find the dispersion relation.

(b) Solve the dispersion relation for $p + C|k|$, where $p = \gamma - i\omega$. Your solution should reduce to the case of counter-streaming delta function beams in the limit $C \rightarrow 0$.

- (c) Show that the solution for $p + C|k|$ in (b) is either purely real or purely imaginary. Which of these cases could possibly lead to instability? What is the frequency ω in this case?
- (d) Show that the condition for instability is $U > C$. Hint: It is useful to know that the fastest growth rate occurs in the vicinity of $k = 0$. Prove this fact by expanding your solution to (b) in powers of k and examining the solution for the growth rate γ in the vicinity of $k = 0$.
- (e) Sketch the shape of $F_0(v_z)$ when $U = C$.

8.17. A plasma consists of two electron beams with a reduced one-dimensional distribution function given by a counter-streaming Cauchy distribution

$$F_0(v_z) = \frac{C}{2\pi} \left[\frac{1}{C^2 + (v_z - U)^2} + \frac{1}{C^2 + (v_z + U)^2} \right].$$

- (a) Show that the distribution function is double humped only if $C < \sqrt{3}U$.
 (b) Using the Penrose criterion, show that instability occurs only if $C < U$.

8.18. Following the procedure developed by Nyquist, make a plot of the mapping of the imaginary p -axis onto the complex D -plane for three cases, $C < U$, $C = U$, and $C > U$, in Problem 8.17.

8.19. A plasma consists of a Maxwellian distribution of ions at temperature T_i and electrons at temperature T_e . The average velocity of the ions is zero and the electrons are drifting with an average velocity U_e .

- (a) Make an appropriate modification of the derivation involving the plasma dispersion function to take into account the drifting electrons.
 (b) Using the low-phase velocity approximation for the electrons and the high-phase velocity approximation for the ions, obtain an approximate expression for the normalized frequencies x_i and y_i of the ion acoustic mode. Assume that $k\lambda_{De} \ll 1$.
 (c) Show that the growth rate is approximately

$$\frac{\gamma}{\omega} = -\sqrt{\frac{\pi}{8}} \left\{ \sqrt{\frac{m_e}{m_i}} \left(1 - \frac{U_e}{C_A} \right) + \left(\frac{T_e}{T_i} \right)^{3/2} \exp\left(-\frac{T_e}{2T_i}\right) \right\}$$

where

$$C_A = \sqrt{\frac{\kappa T_e}{m_i}}$$

is the ion acoustic speed.

- (d) Show that the instability condition for the current-driven ion acoustic mode is

$$|U_e| > \sqrt{\frac{\kappa T_e}{m_i}} \left(1 + \sqrt{\frac{m_i}{m_e}} \left(\frac{T_e}{T_i} \right)^{3/2} \exp\left(-\frac{T_e}{2T_i}\right) \right).$$

- (e) Plot $|U_e|$ as a function of T_e/T_i . Where does the approximation fail?

References

- Arfken, G. 1970. *Mathematical Methods for Physicists*. Academic Press: New York, pp. 311–315.
- Armstrong, T. P. 1967. Numerical studies of the nonlinear Vlasov equation. *Phys. Fluids* **10**, 1269–1280.
- Buneman, O. 1959. Dissipation of currents in ionized media. *Phys. Rev.* **115**, 503–517.
- Dennery, P., and Krzywinski, A. 1967. *Mathematics for Physicists*. Harper and Row: New York, p. 64.
- Flanigan, F. J. 1972. *Complex Variables: Harmonic and Analytic Functions*. Dover Publications (1983 edition), pp. 272–275.
- Fried, B. D., and Conte, S. D. 1961. *The Plasma Dispersion Function*. Academic Press: New York, pp. 2–3.
- Gardner, C. S. 1963. Bound on the energy available from a plasma. *Phys. Fluids* **6**, 839–840.
- Ginzburg, V. L., and Zhelezniakov, V. V. 1958. On the possible mechanism of sporadic solar radio emission (radiation in an isotropic plasma). *Sov. Astron. AJ* **2**, 653–666.
- Gurnett, D. A., Hospodarsky, G. B., Kurth, W. S., Williams, D. J., and Bolton, S. J. 1993. Fine structure of Langmuir waves produced by a solar electron event. *J. Geophys. Res.* **98**, 5631–5637.
- Landau, L. 1946. On the vibration of the electron plasma. *J. Phys. (USSR)* **X** (1), 85–94.
- Malmberg, J. H., and Wharton, C. B. 1966. Dispersion of electron plasma waves. *Phys. Rev. Lett.* **17**, 175–178.
- Nicholson, D. R. 1983. *Introduction to Plasma Theory*. Krieger Publishing: Malabar, FL, pp. 87–96.
- Nyquist, H. 1932. Regeneration theory. *Bell System Tech. J.* **11**, 126–147.
- O’Neil, T. M. 1965. Collisionless damping of nonlinear plasma oscillations. *Phys. Fluids* **8**, 2255–2262.
- Penrose, O. 1960. Electrostatic instabilities of a uniform non-Maxwellian plasma. *Phys. Fluids* **3**, 258–265.
- Vlasov, A. A. 1945. On the kinetic theory of an assembly of particles with collective interaction. *J. Phys. (USSR)* **9**, 25–44.

Further reading

- Chen, F. F. *Introduction to Plasma Physics and Controlled Fusion*. Plenum Press: New York, 1990, Chapter 7.

- Nicholson, D. R. *Introduction to Plasma Theory*. Krieger Publishing: Malabar, FL, 1983, Chapter 6.
- Stix, T. H. *Waves in Plasmas*. American Institute of Physics: New York, 1992, Chapter 8.
- Swanson, D. G. *Plasma Waves*. Academic Press: New York, 1989, Section 4.2.

In this chapter we discuss the propagation of small amplitude waves in a hot magnetized plasma. Just as for a cold plasma, the presence of a static zero-order magnetic field in a hot plasma leads to a wide variety of new phenomena. Because the zero-order motions of the particles in a magnetized plasma consist of circular orbits around the magnetic field, some type of resonance can be expected when the wave frequency is equal to the cyclotron frequency. In a cold plasma, this resonance is the same for all particles of a given charge-to-mass ratio, and gives rise to the well-defined cyclotron resonances described in Chapter 4. In a hot plasma, the frequency “felt” by a particle is Doppler shifted by the thermal motion of the particle along the static magnetic field. For a given parallel velocity, resonance occurs when the frequency in the guiding center frame of reference of the particle is at the cyclotron frequency, i.e., $\omega' = \omega - k_{\parallel}v_{\parallel} = \omega_c$. Because of the thermal spread in the particle velocities, the resonance is no longer sharp, as it was in a cold plasma, but is now broadened by the thermal motion. The resonant interaction also produces damping, called cyclotron damping, in a manner somewhat analogous to Landau damping. If the cyclotron radius of the particle is a significant fraction of the wavelength, the phase shift introduced by the periodic cyclotron motion of the particles back and forth along the perpendicular component of the wave vector produces a phase modulation at the cyclotron frequency. As we will show, this phase modulation leads to a series of resonances at harmonics of the cyclotron frequency. The general cyclotron resonance condition then becomes $\omega' = \omega - k_{\parallel}v_{\parallel} = n\omega_c$, where n is an integer.

Given the complexity of the cyclotron resonance interactions described above, it should come as no surprise that the mathematical analysis is difficult. Nevertheless, it is possible to derive a general dispersion relation for the propagation of small amplitude waves in a hot magnetized plasma. Because of the complexity of the analysis, we will restrict our discussion to certain special cases that are of general interest.

9.1 Linearization of the Vlasov equation

Before proceeding with the derivation of the dispersion relation, it is useful to discuss some general features of the linearized Vlasov equation when a zero-order magnetic field is present. Since the wave amplitudes are assumed to be small, the Vlasov equation is linearized in the usual way by assuming that the velocity distribution function consists of a constant homogeneous zero-order term, $f_{s0}(\mathbf{v})$, plus a small perturbation, $f_{s1}(\mathbf{v})$, so that

$$f_s(\mathbf{v}) = f_{s0}(\mathbf{v}) + f_{s1}(\mathbf{v}). \quad (9.1.1)$$

Since the plasma is magnetized, we must now assume that the magnetic field consists of a constant homogeneous zero-order term plus a small first-order perturbation, i.e., $\mathbf{B} = \mathbf{B}_0 + \mathbf{B}_1$. For the electric field, it is sufficient to assume that the zero-order term is zero, so that $\mathbf{E} = \mathbf{E}_1$. This assumption is justified on the grounds that, except for exceedingly strong electric fields ($E_0 \geq cB_0$), it is always possible to transform to a frame of reference in which $\mathbf{E}_0 = 0$. Using these definitions it is easy to show that the zero- and first-order linearized expansions of the Vlasov equation (??) are given by

$$\mathbf{v} \times \mathbf{B}_0 \cdot \nabla_{\mathbf{v}} f_{s0} = 0, \quad (9.1.2)$$

and

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \frac{e_m}{m_s} (\mathbf{v} \times \mathbf{B}_0) \cdot \nabla_{\mathbf{v}} f_s + \frac{e_s}{m_s} [\mathbf{E} + \mathbf{v} \times \mathbf{B}] \cdot \nabla_{\mathbf{v}} f_{s0} = 0, \quad (9.1.3)$$

where as usual we have dropped the 1 subscript on the first-order terms. Because of the anticipated azimuthal symmetry with respect to the zero-order magnetic field, it is useful to introduce cylindrical velocity space coordinates $(v_{\perp}, \phi, v_{\parallel})$, with \mathbf{B}_0 along the z axis as shown in Figure 9.1. To understand the meaning of the condition imposed on the zero-order distribution function by equation (9.1.2), we start by writing this equation in rectangular (x, y, z) coordinates

$$(\mathbf{v} \times \mathbf{B}_0) \cdot \nabla_{\mathbf{v}} f_{s0} = \left(v_y \frac{\partial f_{s0}}{\partial v_x} - v_x \frac{\partial f_{s0}}{\partial v_y} \right) B_0 = 0. \quad (9.1.4)$$

Next, we transform this equation to cylindrical velocity space coordinates by using the transformation equations $v_x = v_{\perp} \cos \phi$, $v_y = v_{\perp} \sin \phi$, and $v_z = v_{\parallel}$, which when substituted into the above equation give

$$(\mathbf{v} \times \mathbf{B}_0) \cdot \nabla_{\mathbf{v}} f_{s0} = \left(v_{\perp} \sin \phi \frac{\partial f_{s0}}{\partial v_x} - v_{\perp} \cos \phi \frac{\partial f_{s0}}{\partial v_y} \right) B_0 = 0. \quad (9.1.5)$$

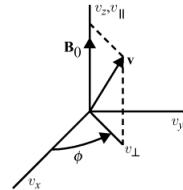


Fig. 9.1 Cylindrical velocity space coordinates, v_{\perp} , v_{\parallel} , and ϕ .

Note from the transformation equations that

$$\frac{\partial v_x}{\partial \phi} = -v_{\perp} \sin \phi \quad \text{and} \quad \frac{\partial v_y}{\partial \phi} = v_{\perp} \cos \phi. \quad (9.1.6)$$

Solving these equations for $\sin \phi$ and $\cos \phi$, and substituting them into equation (9.1.5) then gives

$$(\mathbf{v} \times \mathbf{B}_0) \cdot \nabla_{\mathbf{v}} f_{s0} = - \left[\frac{\partial v_x}{\partial \phi} \frac{\partial f_{s0}}{\partial v_x} + \frac{\partial v_y}{\partial \phi} \frac{\partial f_{s0}}{\partial v_y} \right] B_0 = 0. \quad (9.1.7)$$

Finally, using the chain rule to differentiate $f_{s0}(v_x, v_y)$ with respect to ϕ , one can see that the quantity in the large brackets on the right-hand side is simply $\partial f_{s0}/\partial \phi$. Equation (9.1.2) then becomes

$$(\mathbf{v} \times \mathbf{B}_0) \cdot \nabla_{\mathbf{v}} f_{s0} = -B_0 \frac{\partial f_{s0}}{\partial \phi} = 0. \quad (9.1.8)$$

Since \mathbf{B}_0 is assumed to be non-zero, it follows that $\partial f_{s0}/\partial \phi = 0$. This condition is called the gyrotropic condition, and shows that the zero-order distribution function must be azimuthally symmetric with respect to the magnetic field. This symmetry arises because the particles move in uniform circular motions around the magnetic field. It should be noted that the gyrotropic condition is satisfied only if there is no $\mathbf{E}_0 \times \mathbf{B}_0$ drift (i.e., only if $\mathbf{E}_0 = 0$).

Returning to the first-order linearized Vlasov equation (9.1.3), we note that the third term in this equation has exactly the same form as equation (9.1.8), except f_{s0} is replaced by f_s . Making this substitution and changing $e_s B_0 / m_s$ to ω_{cs} , equation (9.1.3) becomes

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s - \omega_{cs} \frac{\partial f_s}{\partial \phi} + \frac{e_s}{m_s} [\mathbf{E} = \mathbf{v} \times \mathbf{B}] \cdot \nabla_{\mathbf{v}} f_{s0} = 0. \quad (9.1.9)$$

Following the same procedure as in the previous chapter, the next step is to Fourier transform this equation in space ($\nabla \rightarrow i\mathbf{k}$), and in time ($\partial/\partial t \rightarrow -i\omega$), which gives

$$(-i\omega + i\mathbf{k} \cdot \mathbf{v}) \tilde{f}_s - \omega_{cs} \frac{\partial \tilde{f}_s}{\partial \phi} = \frac{e_s}{m_s} [\tilde{\mathbf{E}} + \mathbf{v} \times \tilde{\mathbf{B}}] \cdot \nabla_{\mathbf{v}} f_{s0} = 0. \quad (9.1.10)$$

At a later stage in the analysis, we will convert to Laplace transform notation by making the substitution $\omega \rightarrow ip$. In order to reduce the number of field variables, the first-order magnetic field can be eliminated by using Faraday's law, which when Fourier transformed can be written $\tilde{\mathbf{B}} = \mathbf{k} \times \tilde{\mathbf{E}}/\omega$. Substituting $\tilde{\mathbf{B}} = \mathbf{k} \times \tilde{\mathbf{E}}/\omega$ into the above equation then yields

$$(-i\omega + i\mathbf{k} \cdot \mathbf{v}) \tilde{f}_s - \omega_{cs} \frac{\partial \tilde{f}_s}{\partial \phi} = -\frac{e_s}{m_s} \left[\tilde{\mathbf{E}} + \mathbf{v} \times \left(\frac{\mathbf{k} \times \tilde{\mathbf{E}}}{\omega} \right) \right] \cdot \nabla_{\mathbf{v}} f_{s0}, \quad (9.1.11)$$

where we have organized the terms involving \tilde{f}_s on the left-hand side and the terms involving the electric field on the right-hand side. Since the plasma is gyrotropic, without loss of generality we can assume that \mathbf{k} lies in the x, z plane with components k_{\parallel} and k_{\perp} , parallel and perpendicular to the static magnetic field, as shown in Figure 9.2. From the geometry it is easy to see that

$$\mathbf{k} \cdot \mathbf{v} = k_{\parallel} v_{\parallel} + k_{\perp} v_{\perp} \cos \phi. \quad (9.1.12)$$

After substituting this result into equation (9.1.11) and dividing by $-\omega_{cs}$, the first-order linearized Vlasov equation can be written

$$\frac{\partial \tilde{f}_s}{\partial \phi} - i(\alpha_s + \beta_s \cos \phi) \tilde{f}_s = \frac{e}{m_s \omega_{cs}} \left[\tilde{\mathbf{E}} + \mathbf{v} \times \left(\frac{\mathbf{k} \times \tilde{\mathbf{E}}}{\omega} \right) \right] \cdot \nabla_{\mathbf{v}} f_{s0}, \quad (9.1.13)$$

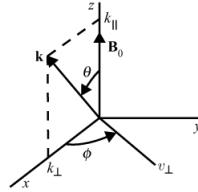


Fig. 9.2 The wave vector, \mathbf{k} , is assumed to lie in the x, z plane at an angle θ with respect to the magnetic field, \mathbf{B}_0 . The parallel and perpendicular components of \mathbf{k} are k_{\parallel} and k_{\perp} .

where to simplify the notation we have introduced the definitions

$$\alpha_s = \frac{k_{\parallel} v_{\parallel} - \omega}{\omega_{cs}} \quad \text{and} \quad \beta_s = \frac{k_{\perp} v_{\perp}}{\omega_{cs}}. \quad (9.1.14)$$

For any given first-order electric field, $\tilde{\mathbf{E}}$, this linear differential equation can be solved to give the first-order perturbation, \tilde{f}_s , in the velocity distribution function.

9.2 Electrostatic waves

Since the analysis of electromagnetic waves is more difficult, we first turn our attention to electrostatic waves. The resulting dispersion relation is called the Harris dispersion relation.

9.2.1 The Harris dispersion relation

The advantage of restricting our attention to electrostatic waves is that the electric field can be written as the gradient of a potential, $\mathbf{E} = -\nabla\Phi$, which when Fourier transformed becomes $\tilde{\mathbf{E}} = -i\mathbf{k}\tilde{\Phi}$. For electrostatic waves one also has $\nabla \times \mathbf{E} = 0$ which, when Fourier transformed, becomes $\mathbf{k} \times \tilde{\mathbf{E}} = 0$. The term $\mathbf{v} \times (\mathbf{k} \times \tilde{\mathbf{E}})$ in equation (9.1.13) is then zero, which provides considerable simplification. After substituting $\tilde{\mathbf{E}} = -i\mathbf{k}\tilde{\Phi}$ into equation (9.1.13), the first-order Vlasov equation becomes

$$\frac{\partial \tilde{f}_s}{\partial \phi} - i(\alpha_s + \beta_s \cos \phi) \tilde{f}_s = \frac{-ie_s}{m_s \omega_{cs}} \tilde{\Phi} \mathbf{k} \cdot \nabla_{\mathbf{v}} f_{s0}. \quad (9.2.1)$$

The above equation is a first-order linear differential equation of the form

$$\frac{df}{dx} = P(x)f = Q(x), \quad (9.2.2)$$

which has a general solution

$$f = e^{-\int^x P(x') dx'} \left[\int^x Q(x') e^{\int^{x'} P(x'') dx''} dx' \right], \quad (9.2.3)$$

where the term $\int P(x) dx$ is called the integrating factor [Conkwright, 1934]. The integrating factor in this case is

$$\begin{aligned} \int^\phi P(\phi') d\phi' &= -i \int^\phi (\alpha_s + \beta_s \cos \phi') d\phi' \\ &= -i(\alpha_s \phi + \beta_s \sin \phi). \end{aligned} \quad (9.2.4)$$

The general solution of equation (9.2.1) is then given by

$$\tilde{f}_s = \frac{-ie_s n_s \tilde{\Phi}}{m_s \omega_c s} e^{i(\alpha_s \phi + \beta_s \sin \phi)} \int^\phi \mathbf{k} \cdot \nabla_{\mathbf{v}} F_{s0} e^{-i(\alpha_s \phi' + \beta_s \sin \phi')} d\phi', \quad (9.2.5)$$

where we have introduced the normalized zero-order distribution function, F_{s0} , via the definition $f_{s0} = n_{s0} F_{s0}$. Following the same procedure used in a hot unmagnetized plasma, the next step is to compute the charge density

$$\tilde{\rho}_q = \sum_s e_s \int_{-\infty}^{\infty} \tilde{f}_s d^3 v, \quad (9.2.6)$$

and substitute the charge density into Poisson's equation, $\nabla^2 \Phi = -\rho_q/\epsilon_0$ which, after Fourier transforming, becomes

$$k^2 \tilde{\Phi} = \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} \tilde{f}_s d^3 v. \quad (9.2.7)$$

The dispersion relation can then be obtained by substituting \tilde{f}_s from equation (9.2.5) into the integral, canceling $\tilde{\Phi}$ which gives on both sides of the equation, and dividing by k^2 , which gives

$$\begin{aligned} D(k, \omega) &= 1 + \sum_s \frac{\omega_{ps}^2}{k^2 \omega_c s} \int e^{i(\alpha_s \phi + \beta_s \sin \phi)} \\ &\times \int^\phi i \mathbf{k} \cdot \nabla_{\mathbf{v}} F_{s0} e^{-i(\alpha_s \phi' + \beta_s \sin \phi')} d\phi' d\phi v_\perp dv_\perp dv_\parallel = 0, \end{aligned} \quad (9.2.8)$$

where the velocity space volume element is written as $d^3 v = d\phi v_\perp dv_\perp dv_\parallel$. Unfortunately, the above equation is not yet in a useful form. To simplify the equation, we proceed by

writing $\nabla_v F_{s0}$ in cylindrical coordinates, noting from equation (9.1.8) that $\partial F_{s0}/\partial\phi = 0$, and taking the dot-product with $i\mathbf{k}$. The first term in the integrand is then given by

$$i\mathbf{k} \cdot \nabla_v F_{s0} = ik_{\parallel} \frac{\partial F_{s0}}{\partial v_{\parallel}} + ik_{\perp} \frac{\partial F_{s0}}{\partial v_{\perp}} \cos\phi'. \quad (9.2.9)$$

By writing $\cos\phi'$ in terms of complex exponentials, the $d\phi'$ integral in equation (9.2.8) can be written

$$\begin{aligned} \int^{\phi} \dots d\phi' &= ik_{\parallel} \frac{\partial F_{s0}}{\partial v_{\parallel}} \int^{\phi} e^{-i(\alpha_s \phi' + \beta_s \sin \phi')} d\phi' + ik_{\perp} \frac{\partial F_{s0}}{\partial v_{\perp}} \\ &\times \int^{\phi} \frac{1}{2} (e^{i\phi'} + e^{-i\phi'}) e^{-i(\alpha_s \phi' + \beta_s \sin \phi')} d\phi' \end{aligned} \quad (9.2.10)$$

The integrals in the above equation can be evaluated by using the following expansion [from Arfken, 1970]

$$e^{-i\beta_s \sin \phi'} = \sum_{n=-\infty}^{\infty} J_n(\beta_s) e^{-in\phi'}, \quad (9.2.11)$$

where $J_n(\beta_s)$ is the n th order Bessel function. Using this expansion, the first integral on the right-hand side of equation (9.2.10) becomes

$$\begin{aligned} \int^{\phi} e^{-i(\alpha_s \phi' + \beta_s \sin \phi')} d\phi' &= \sum_n J_n(\beta_s) \int^{\phi} e^{-i(\alpha_s + n)\phi'} d\phi' \\ &= i \sum_n \frac{J_n(\beta_s)}{\alpha_s + n} e^{-i(\alpha_s + n)\phi}. \end{aligned} \quad (9.2.12)$$

Similarly, the second integral on the right-hand side of equation (9.2.10) becomes

$$\begin{aligned} &\int^{\phi} (e^{i\phi'} + e^{-i\phi'}) e^{-i(\alpha_s \phi' + \beta_s \sin \phi')} d\phi' \\ &= \sum_n J_n(\beta_s) \int^{\phi} [e^{-i(\alpha_s - 1 + n)\phi'} + e^{-i(\alpha_s + 1 + n)\phi'}] d\phi' \\ &= i \sum_n J_n(\beta_s) \left[\frac{e^{-i(\alpha_s - 1 + n)\phi}}{\alpha_s - 1 + n} + \frac{e^{-i(\alpha_s + 1 + n)\phi}}{\alpha_s + 1 + n} \right]. \end{aligned} \quad (9.2.13)$$

Next, multiplying equation (9.2.10) by the term

$$e^{i(\alpha_s \phi + \beta_s \sin \phi)} = \sum_m J_m(\beta_s) e^{i(\alpha_s + m)\phi} \quad (9.2.14)$$

gives

$$\begin{aligned} e^{i(\alpha_s \phi + \beta_s \sin \phi)} \int^\phi \dots d\phi' = & -k_{\parallel} \frac{\partial F_{s0}}{\partial v_{\parallel}} \sum_{n,m} J_m J_n \left[\frac{e^{i(m-n)\phi}}{\alpha_s + n} \right] \\ & - \frac{1}{2} k_{\perp} \frac{\partial F_{s0}}{\partial v_{\perp}} \sum_{n,m} J_m J_n \left[\frac{e^{i(m-n+1)\phi}}{\alpha_s + n - 1} + \frac{e^{i(m-n-1)\phi}}{\alpha_s + n + 1} \right]. \end{aligned} \quad (9.2.15)$$

In the second summation the index can be relabeled to give

$$\sum_{n,m} \frac{J_m J_{n+1}}{\alpha_s + n} e^{i(m-n)\phi} + \frac{J_m J_{n-1}}{\alpha_s + n} e^{i(m-n)\phi} = \sum_{n,m} J_m [J_{n+1} + J_{n-1}] \frac{e^{i(m-n)\phi}}{\alpha_s + n}. \quad (9.2.16)$$

Using the Bessel function recursion formula $J_{n+1} + J_{n-1} = (2n/\beta_s)J_n$, this sum can be written in the more compact form

$$\sum_{m,n} \frac{2n J_m J_n}{\beta_s (\alpha_s + n)} e^{i(m-n)\phi}. \quad (9.2.17)$$

Combining all the terms then gives

$$e^{i(\alpha_s \phi + \beta_s \sin \phi)} \int^\phi \dots d\phi' = \sum_{n,m} \frac{J_m J_n}{\alpha_s + n} \left[k_{\parallel} \frac{\partial F_{s0}}{\partial v_{\parallel}} + \frac{n\omega_{cs}}{v_{\perp}} \frac{\partial F_{s0}}{\partial v_{\perp}} \right] e^{i(m-n)\phi}, \quad (9.2.18)$$

where in the second term on the right we have eliminated β_s by using $\beta_s = k_{\perp} v_{\perp} / \omega_{cs}$ from equation (9.1.14). Inserting this result into equation (9.2.8), we obtain

$$\begin{aligned} D(k, \omega) = & 1 - \sum_s \frac{\omega_{ps}^2}{k^2 \omega_{cs}} \sum_{n,m} \int_{-\infty}^{\infty} \int_0^{\infty} v_{\perp} dv_{\perp} dv_{\parallel} \frac{J_m J_n}{\alpha_s + n} \\ & \times \left[k_{\parallel} \frac{\partial F_{s0}}{\partial v_{\parallel}} + \frac{n\omega_{cs}}{v_{\perp}} \frac{\partial F_{s0}}{\partial v_{\perp}} \right] \int_0^{2\pi} e^{i(m-n)\phi} d\phi = 0. \end{aligned} \quad (9.2.19)$$

The ϕ integral in the above equation is zero unless $m = n$, in which case it is 2π . Finally, reintroducing the definitions for α_s and β_s from equation (9.1.14), the dispersion relation can be written in the form

$$\begin{aligned} D(k, \omega) = & 1 - \sum_s \frac{\omega_{ps}^2}{k^2} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{J_n^2 (k_{\perp} v_{\perp} / \omega_{cs})}{k_{\parallel} v_{\parallel} - \omega + n\omega_{cs}} \left[k_{\parallel} \frac{\partial F_{s0}}{\partial v_{\parallel}} + \frac{n\omega_{cs}}{v_{\perp}} \frac{\partial F_{s0}}{\partial v_{\perp}} \right] \\ & \times 2\pi v_{\perp} dv_{\perp} dv_{\parallel} = 0. \end{aligned} \quad (9.2.20)$$

This equation is called the Harris dispersion relation after Harris [1959], who first derived this result.

The Harris dispersion relation gives a very general result for small amplitude electrostatic waves propagating in a hot magnetized plasma and is valid for any direction of propagation and any choice of plasma parameters. The origin of the cyclotron harmonics, $n\omega_{cs}$,

in the dispersion relation can be traced to the term $\exp[i\beta_s \sin \phi]$ and has the following interpretation. Note that the parameter $\beta_s = k_\perp v_\perp / \omega_{cs}$ is the product of the cyclotron radius, $\rho_c = v_\perp / \omega_c$, and the perpendicular wave number, k_\perp . If the wavelength is comparable to the cyclotron radius, so that $k_\perp \rho_c \sim 1$, then a substantial phase shift is introduced in the applied electric field by the zero-order cyclotron motion of the particle. The origin of the phase shift is illustrated in Figure 9.3. To understand the significance of the phase shift, consider an applied electric field of the form $E_0 \exp[i(k_\perp x - \omega t)]$. If the zero-order motion along the x axis is $x = \rho_c \sin \omega_c t$, the electric field at the position of the particle is given by

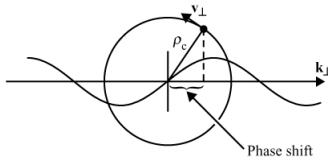


Fig. 9.3 For a finite cyclotron radius, ρ_c , a phase shift is introduced by the cyclotron motion of the particle around the magnetic field. This phase shift is responsible for the resonances at $n\omega_c$.

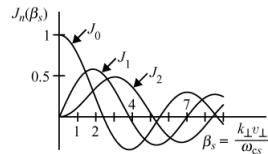


Fig. 9.4 Plots of the zero-order, first-order, and second-order Bessel functions, $J_0(\beta_s)$, $J_1(\beta_s)$, and $J_2(\beta_s)$.

$$E = E_0 e^{i\beta_s \sin \omega_c t} e^{-i\omega t} = E_0 \sum_{n=-\infty}^{\infty} J_n(\beta_s) e^{i(n\omega_c - \omega)t}, \quad (9.2.21)$$

where we have used the harmonic expansion given by equation (9.2.11) for the $\exp(i\beta_s \sin \omega_c t)$ term. Since the phase shift term, $\beta_s \sin \omega_c t$, is periodic with angular frequency ω_c , the above equation shows that the electric field force can be expressed as an expansion in harmonics of the cyclotron frequency, ω_c . The coefficients of this expansion are the Bessel functions $J_n(\beta_s)$. When $\omega \simeq n\omega_c$ the electric field force “felt” by the particle has a component in phase with the motion of the particle and a strong resonant interaction occurs. The Bessel function term gives the “strength” of this interaction, and is controlled by β_s . The functional forms of the first three Bessel functions are shown in Figure 9.4. It can be shown that as β_s goes to zero, $J_0 \rightarrow 1$, $J_1 \rightarrow \beta_s/2$, and $J_2 \rightarrow \beta_s^2/8$. Also, $J_n(-\beta_s) = (-1)^n J_n(\beta_s)$ and $J_{-n}(\beta_s) = (-1)^n J_n(\beta_s)$, see Watson [1995].

9.2.2 The low-temperature, long wavelength limit

Before discussing the general solution of the Harris dispersion relation, it is useful to consider the low-temperature, long wavelength limit. If the perpendicular temperature is sufficiently low ($v_{\perp} \rightarrow 0$) and/or the perpendicular wavelength is sufficiently long ($k_{\perp} \rightarrow 0$), the parameter $\beta_s = k_{\perp}v_{\perp}/\omega_{cs}$ can be regarded as a small quantity (i.e., $\beta_s \ll 1$). In this limit the Bessel functions in (9.2.11) can be approximated to first order in β_s by three terms, $J_0 \approx 1$, $J_1 \approx \beta_s/2$, and $J_{-1} \approx -\beta_s/2$. All the remaining Bessel function terms are of higher order in β_s and can be omitted. Substituting these approximations into the dispersion relation, keeping only the $n = 0$ term associated with $\partial F_{s0}/\partial v_{\parallel}$ and the $n = \pm 1$ term associated with $\partial F_{s0}/\partial v_{\perp}$, and expanding the v_{\parallel} integral in the small parameters $k_{\parallel}v_{\parallel}/\omega$ and $k_{\parallel}v_{\parallel}/(\omega \pm \omega_c)$, the integral can be written as a series of integrals involving moments in v_{\parallel} , very similar to the high-phase velocity expansion in equation (??). Keeping only the zero-order moments (i.e., the zero temperature limit), it is easy to show that the Harris dispersion relation simplifies to

$$D_0(k, \omega) = \left[1 - \sum_s \frac{\omega_{ps}^2}{\omega^2} \right] \cos^2 \theta + \left[1 - \sum_s \frac{\omega_{ps}^2}{\omega^2 - \omega_{cs}^2} \right] \sin^2 \theta = 0, \quad (9.2.22)$$

where $\cos \theta = k_{\parallel}/k$ and $\sin \theta = k_{\perp}/k$. The terms in the brackets can be immediately recognized as the functions P and S in cold plasma theory, i.e., see equations (??) and (??). The dispersion relation can then be written

$$D_0(k, \omega) = P \cos^2 \theta + S \sin^2 \theta = 0. \quad (9.2.23)$$

For propagation parallel to the magnetic field, $\theta = 0$, the above equation reduces to $P = 0$, which gives the electron plasma oscillations encountered in cold plasma theory. For propagation perpendicular to the magnetic field, $\theta = \pi/2$, the dispersion relation reduces to $S = 0$, which gives the hybrid resonances. At intermediate angles the dispersion relation corresponds to resonance cones at $\tan^2 \theta = -P/S$, i.e., equation (??). Thus, the low-temperature, long wavelength limit of the Harris dispersion relation reduces to the electrostatic waves encountered in cold plasma theory.

9.2.3 The Bernstein modes

To explore further the solutions of the Harris dispersion relation, we next consider the special case of an isotropic Maxwellian velocity distribution with the wave vector perpendicular to the magnetic field (i.e., $k_{\parallel} = 0$). To simplify the analysis, the ions are assumed to be immobile, so that only the electron motions need be considered. Keeping only the electron terms, the Harris dispersion relation (9.2.20), with $k_{\parallel} = 0$, $\omega_c = |\omega_{ce}| = -\omega_{ce}$, and $\omega_{pe} = \omega_p$, can be written

$$\begin{aligned} D(k_{\perp}, \omega) = & 1 - \frac{\omega_p^2}{k_{\perp}^2} \sum_{n=-\infty}^{\infty} \frac{n\omega_c}{-\omega + n\omega_c} \\ & \times \int_{-\infty}^{\infty} \int_0^{\infty} J_n^2 \left(\frac{k_{\perp} v_{\perp}}{\omega_c} \right) \frac{\partial F_0}{\partial v_{\perp}} 2\pi dv_{\perp} dv_{\parallel} = 0, \end{aligned} \quad (9.2.24)$$

where we have changed n to $-n$ and used the fact that $J_{-n}^2(-\beta_s) = J_n^2(\beta_s)$. A further simplification is made by combining the positive ($+n$) and negative ($-n$) terms

$$\frac{n\omega_c}{-\omega + n\omega_c} + \frac{(-n)\omega_c}{-\omega + (-n)\omega_c} = \frac{-2n^2\omega_c^2}{\omega^2 - n^2\omega_c^2}, \quad (9.2.25)$$

which gives

$$\begin{aligned} D(k_{\perp}, \omega) = & 1 + \frac{2\omega_p^2}{k_{\perp}^2} \sum_{n=1}^{\infty} \frac{n^2\omega_c^2}{\omega^2 - n^2\omega_c^2} \\ & \times \int_{-\infty}^{\infty} \int_0^{\infty} J_n^2 \left(\frac{k_{\perp} v_{\perp}}{\omega_c} \right) \frac{\partial F_0}{\partial v_{\perp}} 2\pi dv_{\perp} dv_{\parallel} = 0, \end{aligned} \quad (9.2.26)$$

where the summation now extends only over the positive integers. In contrast to the situation for a Maxwellian velocity distribution function in an unmagnetized plasma, the integral in the dispersion relation can be performed in closed form. This is made possible by the following identity [from Watson, 1995]

$$\int_0^{\infty} J_n^2(au) e^{-u^2} u du = \frac{1}{2} \exp\left(-\frac{a^2}{2}\right) I_n\left(\frac{a^2}{2}\right), \quad (9.2.27)$$

where $I_n(x)$ is the modified Bessel function of order n . These functions can be computed using the following power series expansion

$$I_n(x) = \sum_{m=0}^{\infty} \frac{1}{m!(m+|n|)!} \left(\frac{x}{2}\right)^{2m+|n|}. \quad (9.2.28)$$

The modified Bessel functions increase monotonically with increasing x as shown in Figure 9.5. After completing the required integration in equation (9.2.26), the dispersion relation can be written in the following form

$$D(k_{\perp}, \omega) = 1 - \sum_{n=1}^{\infty} \frac{2\omega_p^2}{\beta_c^2 \omega_c^2} \frac{\Gamma_n(\beta_c)}{(\omega/n\omega_c)^2 - 1} = 0, \quad (9.2.29)$$

where the function $\Gamma_n(\beta_c)$ in the above equation is defined by

$$\Gamma_n(\beta_c) = e^{-\beta_c^2} I_n(\beta_c^2), \quad (9.2.30)$$

and β_c is defined by

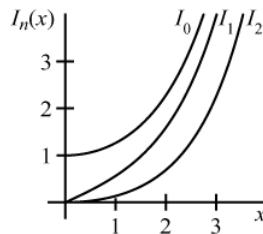


Fig. 9.5 Plots of the zero-order, first-order, and second-order modified Bessel functions, $I_0(x)$, $I_1(x)$ and $I_2(x)$.

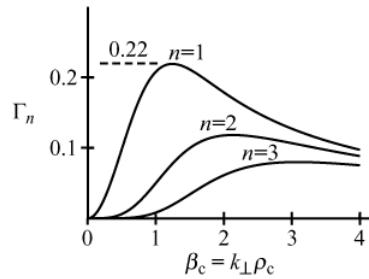


Fig. 9.6 Plots of $\Gamma_n(\beta_c)$ for $n = 1$, 2, and 3.

$$\beta_c = k_{\perp} \frac{\sqrt{\kappa T/m}}{\omega_c}. \quad (9.2.31)$$

Note that the quantity multiplying k_{\perp} in the above equation is simply the thermal cyclotron radius, $\rho_c = \sqrt{\kappa T/m/\omega_c}$, i.e., the cyclotron radius of an electron moving at the thermal velocity, so that $\beta_c = k_{\perp}\rho_c$. Plots of $\Gamma_n(\beta_c)$ as a function of β_c are shown in Figure 9.6 for $n = 1$, 2, and 3. As can be seen, the Γ_n functions all have a single maximum that decreases and shifts to larger β_c values as n increases. For $n = 1$, the peak value is 0.22 at $\beta_c = 1.24$.

The dispersion relation given by equation (9.2.29) was first derived by Bernstein [1958], and the corresponding modes of propagation are called the Bernstein modes. The general nature of the roots of the dispersion relation can be determined by plotting $D(k_{\perp}, \omega)$ as a function of ω for a specific value of $\beta_c = k_{\perp}\rho_c$ and then looking for the points where $D = 0$. Notice that the thermal cyclotron radius, ρ_c , is a basic scale factor in the perpendicular wavelength. Note also that as $k_{\perp}\rho_c \rightarrow 0$, the dispersion relation, $D(k_{\perp}, \omega)$, approaches the cold plasma equation, S , given by equation (??). A representative plot of $D(k_{\perp}, \omega)$ is shown in Figure 9.7. Since $D(k_{\perp}, \omega)$ has a single-order pole at each of the cyclotron harmonics, the roots of the dispersion relation consist of an infinite number of frequencies, $\omega_1, \omega_2, \dots$, with one root between each of the adjacent cyclotron harmonics. By solving the dispersion relation, $D(k_{\perp}, \omega) = 0$, as a function of $\beta_c = k_{\perp}\rho_c$, the frequency of each mode, ω_n , can be determined as a function of $k_{\perp}\rho_c$. This procedure must be done numerically. A typical plot of ω_n as a function of $\beta_c = k_{\perp}\rho_c$ is shown in Figure 9.8. For the parameters used, $\omega_p = 2.5\omega_c$, the ratio of the plasma frequency to the cyclotron frequency is such that the upper hybrid resonance frequency, ω_{UH} , is located between the second and third harmonics

of the electron cyclotron frequency. Note that the branch of the dispersion relation between these two harmonics goes to ω_{UH} as $k_{\perp}\rho_c \rightarrow 0$. This is to be expected from our previous discussion of the low-temperature, long wavelength limit. Also, note that the qualitative form of the dispersion curves changes at ω_{UH} . For frequencies below ω_{UH} , the allowed frequency range extends across the entire band between adjacent harmonics, whereas above ω_{UH} the allowed frequency range is restricted to a narrow frequency range just above the harmonic.

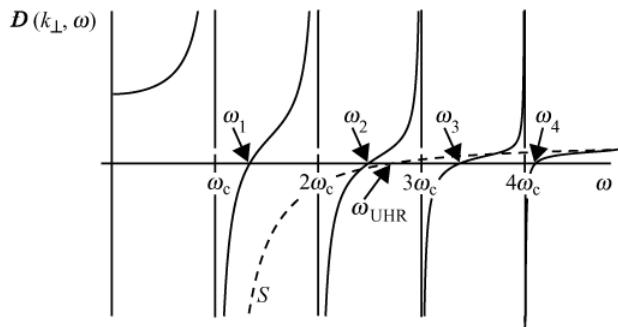


Fig. 9.7 A plot of $D(k_{\perp}, \omega)$ as a function of frequency. The dashed line is the cold plasma limit, S .

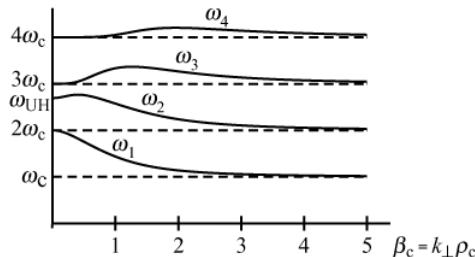


Fig. 9.8 A plot of the solutions of $D(k_{\perp}, \omega) = 0$ as a function of the perpendicular wave number, k_{\perp} . These solutions are called the Bernstein modes.

The existence of electrostatic waves near the harmonics of the cyclotron frequency is a totally new, hot plasma effect not predicted by either cold plasma theory or the more general fluid theory. Even though one might expect some type of cyclotron damping because the plasma is hot, for propagation perpendicular to the magnetic field ($k_{\parallel} = 0$) there is no damping. This occurs because for $k_{\parallel} = 0$ the resonance velocity, $v_{\parallel} = (\omega - n\omega_{cs})/k_{\parallel}$, is infinite, so no particles can resonate with the wave. The absence of damping means that there is no short wavelength limit to the electrostatic waves, contrary to the results found in Chapter 8 for an unmagnetized plasma. However, this situation is a special case applicable only for propagation exactly perpendicular to the magnetic field. For any angle slightly away from perpendicular the damping reappears. Wavelengths less than the Debye length are then strongly damped.

Both laboratory and space plasma measurements have clearly verified the existence of

the Bernstein modes. In space experiments, the clearest examples occur in the records from satellite-borne ionospheric sounders, such as were discussed in Chapter ???. Since the Bernstein modes have no damping, once excited the waves decay very slowly and are easy to detect. The received signal typically consists of a series of “spikes” or “resonances” near harmonics of the electron cyclotron frequency, such as shown in Figure 9.9. Although the sounding pulse excites waves over a broad range of frequencies, the waves detected after the initial transient has decayed are those that have zero group velocity, since only these waves stay in the vicinity of the spacecraft. As can be seen from Figure 9.8, the group velocity, $\partial\omega/\partial k_{\perp}$, goes to zero at (1) the harmonics of the electron cyclotron frequency, $n\omega_c$, (2) the upper hybrid resonance frequency, f_{UH} , and (3) a series of frequencies, f_Q , called the Q resonances. The Q resonances are the frequencies where the dispersion curves in Figure 9.8 reach a maximum (i.e., where $\partial\omega/\partial k_{\perp} = 0$). All three types of resonances can be seen in Figure 9.9.

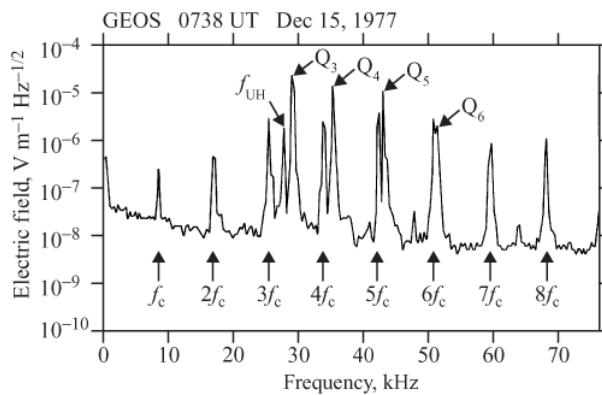


Fig. 9.9 An ionospheric sounder response showing the various resonances associated with the Bernstein modes [data provided by P. Canu].

9.2.4 Instabilities

Next we consider various types of electrostatic instabilities that can occur in a hot magnetized plasma. Two broad classes of instabilities can be identified: (1) anisotropy-driven instabilities, and (2) current-driven instabilities. We will discuss examples of each.

9.2.4.1 Anisotropy-driven instabilities

For propagation perpendicular or nearly perpendicular to the magnetic field, an interesting class of instabilities occurs between the harmonics of the electron cyclotron frequency that are driven by anisotropic electron velocity distributions. To illustrate the existence of such instabilities, we start by considering a ring-type electron velocity distribution, such as shown in Figure 9.10. Distribution functions of this type can be produced by injecting a beam of electrons into the plasma at pitch angles nearly perpendicular to the magnetic

field. It is easy to show that such distributions are unstable. As a simple example, consider the following delta function distribution

$$F_0(v_\perp, v_\parallel) = \frac{1}{2\pi V_0} \delta(v_\parallel) \delta(v_\perp - V_0), \quad (9.2.32)$$

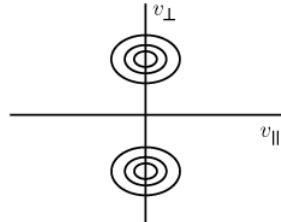


Fig. 9.10 A representative ring-type velocity distribution.

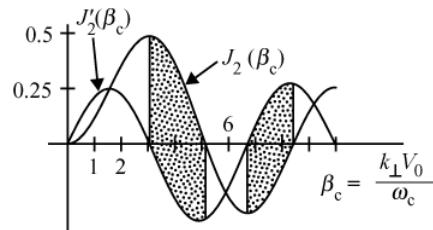


Fig. 9.11 Plots of $J'_2(\beta_c)$ and $J_2(\beta_c)$. Instabilities can occur in the shaded regions where the product $J_n(\beta_c)J'_n(\beta_c)$ is negative. The exact instability condition depends on the ratio ω_p/ω_c .

which represents a narrow ring of particles in velocity space centered at $v_\parallel = 0$ and $v_\perp = V_0$. If we restrict our attention to waves propagating exactly perpendicular to the magnetic field ($k_\parallel = 0$), the dispersion relation can be evaluated in a straightforward way by substituting the above equation into equation (9.2.26) and integrating by parts once, which gives

$$\mathcal{D}(k_\perp, \omega) = 1 - \frac{4\omega_p^2}{k_\perp V_0 \omega_c} \sum_{n=1}^{\infty} \frac{J_n(\beta_c)J'_n(\beta_c)}{(\omega/n\omega_c)^2 - 1} = 0, \quad (9.2.33)$$

where $\beta_c = k_\perp V_0 / \omega_c$ is 2π times the ratio of the cyclotron radius to the wavelength. The signs of the various terms in the summation are determined by the sign of the product $J_n(\beta_c)J'_n(\beta_c)$. Representative plots of $J_n(\beta_c)$ and $J'_n(\beta_c)$ are shown in Figure 9.11 for $n = 2$. For small velocities ($\beta_c \ll 1$), the signs of the $J_n(\beta_c)J'_n(\beta_c)$ terms are all positive, and the solutions to the dispersion relation have the same general form as for the Bernstein modes, with one root between each cyclotron harmonic. However, as the ring velocity V_0 increases, β_c increases until eventually the sign of one of the $J_n(\beta_c)J'_n(\beta_c)$ terms becomes negative, as in the shaded regions of Figure 9.11. The plot of $\mathcal{D}(k_\perp, \omega)$ versus ω then changes form, with the branches adjacent to this cyclotron harmonic curving alternately concave downward and then concave upward. As V_0 is increased further, a point is reached where these

branches no longer cross the $D = 0$ axis, as illustrated in Figure 9.12 for the $n = 2$ term. At this point the two real roots associated with these branches change to a complex conjugate pair, one of which is unstable. Because of the complexity of $D(k_\perp, \omega)$, the exact point at which the instability starts must be evaluated numerically. Detailed numerical studies by Tataronis and Crawford [1970] show that a minimum plasma density exists for an instability to occur, at $\omega_p/\omega_c \approx 4.13$ for the 0 to ω_c band, $\omega_p/\omega_c \approx 2.61$ for the ω_c to $2\omega_c$ band, etc. Tataronis and Crawford also explored a number of other similar velocity distributions, such as a spherical shell distribution and a combination of a Maxwellian and a ring distribution, all of which are unstable for certain choices of the parameters. In many respects these instabilities are like the cold beam instabilities discussed in Chapter ?? for unmagnetized plasmas. The ring distribution essentially produces two oppositely directed beams. The results are modified, of course, by the presence of magnetic field forces.

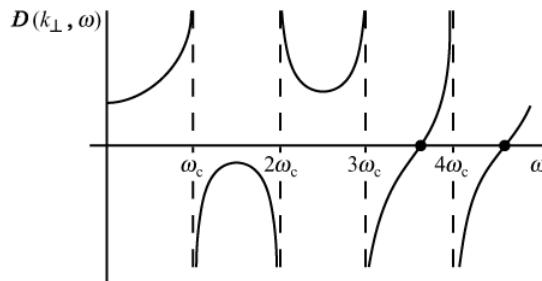


Fig. 9.12 A plot of $D(k_\perp, \omega)$ for an unstable distribution. In this case the instability is associated with the $n = 2$ term, for which the product $J'_2(\beta_c)J_2(\beta_c)$ is negative.

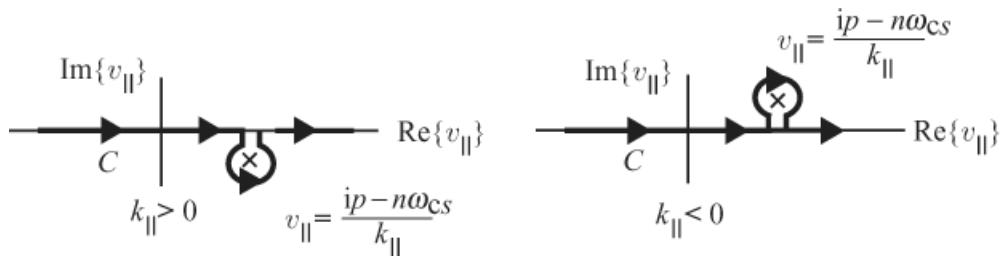


Fig. 9.13 The v_{\parallel} integration contour for the two cases $k_{\parallel} > 0$ and $k_{\parallel} < 0$.

In addition to the ring-type instabilities described above, which are non-resonant fluid-like instabilities, a variety of other instabilities, very similar to the Bernstein modes, can be generated by resonant interactions. To investigate these instabilities, we must convert the Harris dispersion relation (9.2.20) to the Landau formalism. This conversion is accomplished by replacing ω by ip and changing the integration along the v_{\parallel} axis such that it always passes below the pole at $v_{\parallel} = (ip - n\omega_{CS})/k_{\parallel}$ for $k_{\parallel} > 0$, and above the pole for $k_{\parallel} < 0$, as shown in Figure 9.13. Following the approach developed in Chapter ??, we use the weak growth rate approximation, which involves separating the dispersion relation into its real and imaginary parts using the Plemelj relation (see Section ??). Again keeping

only the electron terms (with $\omega_c = -\omega_{ce}$, $\omega_p = \omega_{pe}$, and changing n to $-n$), the real and imaginary parts of the Harris dispersion relation are

$$\begin{aligned} D_r = & 1 - \frac{\omega_p^2}{k^2} \sum_n P \int_{-\infty}^{\infty} \int_0^{\infty} \frac{J_n^2(k_{\perp} v_{\perp}/\omega_c)}{k_{\parallel} v_{\parallel} - \omega + n\omega_c} \\ & \times \left[k_{\parallel} \frac{\partial F_0}{\partial v_{\parallel}} + \frac{n\omega_c}{v_{\perp}} \frac{\partial F_0}{\partial v_{\perp}} \right] 2\pi v_{\perp} dv_{\perp} dv_{\parallel} \end{aligned} \quad (9.2.34)$$

and

$$\begin{aligned} D_i = & -\pi \frac{k_{\parallel}}{|k_{\parallel}|} \frac{\omega_p^2}{k^2} \sum_n \int_0^{\infty} J_n^2 \left(\frac{v_{\perp} k_{\perp}}{\omega_c} \right) \\ & \times \left[k_{\parallel} \frac{\partial F_0}{\partial v_{\parallel}} + \frac{n\omega_c}{v_{\perp}} \frac{\partial F_0}{\partial v_{\perp}} \right] 2\pi v_{\perp} dv_{\perp} \Big|_{v_{\parallel Res} = \frac{\omega - n\omega_c}{k_{\parallel}}} , \end{aligned} \quad (9.2.35)$$

where, as indicated, the integral in D_i must be evaluated at the resonance velocity $v_{\parallel Res} = (\omega - n\omega_c)/k_{\parallel}$. To achieve a finite resonance velocity, the wave vector cannot be exactly perpendicular to the magnetic field, even though in most cases it is nearly perpendicular. The growth rate can then be calculated from the weak growth rate approximation (??)

$$\gamma = \frac{-D_i}{\partial D_r / \partial \omega}, \quad (9.2.36)$$

where ω and k are determined by $D_r(\omega, k) = 0$.

Since $\partial D_r / \partial \omega$ is always positive for the Bernstein-like modes (see Figure 9.7), the sign of the growth rate is controlled by the sign of the integral in equation (9.2.35). Consider for the moment only waves with $k_{\parallel} > 0$. The growth rate is then positive if the sign of the integral is positive. The sign of the integral is controlled by the terms $\partial F_0 / \partial v_{\parallel}$ and $\partial F_0 / \partial v_{\perp}$. Any region of the distribution function where either of these terms is positive is a potential source of instability. These regions are called free energy sources. If the integral over perpendicular velocities turns out to be positive for a particular \mathbf{k} (assuming $k_{\parallel} > 0$), then the plasma is unstable. Note that both k_{\parallel} and k_{\perp} are involved in the integral. Also note that k_{\parallel} plays a controlling role in determining the resonance velocity, $v_{\parallel Res} = (\omega - n\omega_c)/k_{\parallel}$. Our problem is to understand the nature of this control.

In the previous chapter we have already encountered instabilities associated with regions where $\partial F_0 / \partial v_{\parallel}$ is positive. This free energy source is characteristic of a beam. The presence of a magnetic field reveals a new type of free energy source that is associated with regions where $\partial F_0 / \partial v_{\perp}$ is positive. This free energy source is characteristic of an anisotropy. Although many different velocity distributions have an anisotropy, the prototypical distribution function that illustrates this type of free energy source is a loss-cone distribution, such as is shown in Figure 9.14. For an instability to occur, the parallel wave number must be such that the parallel resonance velocity, $v_{\parallel Res}$, indicated by the vertical dashed line, passes through the region where $\partial F_0 / \partial v_{\perp}$ is positive. Both positive and neg-

ative values of $\partial F_0 / \partial v_{\perp}$ occur along this resonance line. As can be seen from equation (9.2.35), the integral over perpendicular velocities is of the form

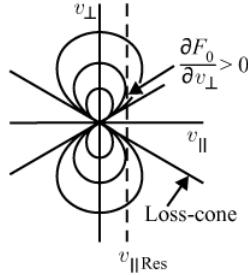


Fig. 9.14 A loss-cone type of distribution function.

$$\int_0^\infty J_n^2\left(\frac{k_{\perp}v_{\perp}}{\omega_c}\right) \frac{\partial F_0}{\partial v_{\perp}} dv_{\perp} \Big|_{v_{\parallel\text{Res}} = \frac{\omega - n\omega_c}{k_{\parallel}}}, \quad (9.2.37)$$

which involves a convolution of the terms $J_n^2(k_{\perp}v_{\perp}/\omega_c)$ and $\partial F_0/\partial v_{\perp}$. For an instability to occur, the integral must be sufficiently positive (assuming $k_{\parallel} > 0$) to over-come any negative contribution from the $k_{\parallel}\partial F_0/\partial v_{\parallel}$ term in equation (9.2.35). A representative plot of the various terms in equation (9.2.37) is shown in Figure 9.15 for a loss-cone distribution. As can be seen from this illustration, the perpendicular

wave number, k_{\perp} , controls the peaks in the Bessel function. The largest positive value for the integral usually occurs when k_{\perp} is such that the first maximum in $J_n^2(k_{\perp}v_{\perp}/\omega_c)$ coincides with the peak in $\partial F_0/\partial v_{\perp}$, which is near the edge of the loss cone. Thus, the Bessel function plays a key role in determining the range of unstable perpendicular wave numbers for each harmonic. We can make a rough estimate of the wave normal angle by noting that, in the unstable region (i.e., where $\partial F_0/\partial v_{\perp} > 0$), one has $k_{\perp}v_{\perp}/\omega_c \sim 1$ and $v_{\parallel} = (\omega - n\omega_c)/k_{\parallel}$. Since v_{\perp}/v_{\parallel} is typically of order one in this region, it follows that $\tan \theta = k_{\perp}/k_{\parallel} \approx \omega_c/(\omega - n\omega_c)$. As with the Bernstein modes, the frequency is usually close to the cyclotron harmonic (see Figure 9.9), especially at the higher harmonics. The wave vector direction is then at a large angle to the magnetic field, typically 70° to 80° .

Electrostatic electron cyclotron instabilities driven by cyclotron resonance interactions near the loss cone are frequently observed in both laboratory and space plasmas. The instabilities usually occur as a series of emission lines between harmonics of the electron cyclotron frequency. These waves are often called $(n + 1/2)f_c$ emissions, even though the frequencies are somewhat less than $(n + 1/2)f_c$. An example showing the spectrum of a series of electrostatic $(n + 1/2)f_c$ emissions detected in Earth's magnetosphere is shown in Figure 9.16. Similar emissions also sometimes occur near the upper hybrid resonance, at frequencies such that $(n + 1/2)f_c \approx f_{UHR}$. For a further discussion of these waves, see Ashour-Abdalla and Kennel [1978], Rönnmark [1983] and Horne [1989].

In addition to the electron cyclotron instabilities, another important class of instabilities occurs near the harmonics of the ion cyclotron frequency. These waves are called electrostatic ion cyclotron waves and are driven by anisotropies in the ion distribution function.

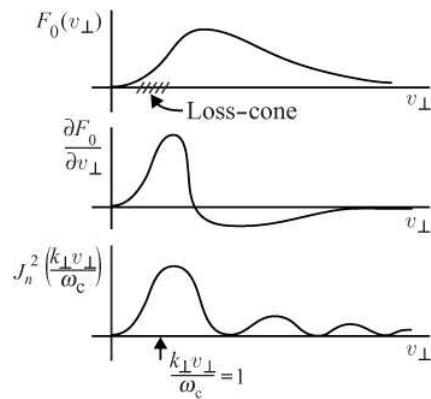


Fig. 9.15 Plots of $F_0(v_\perp)$, $\partial F_0/\partial v_\perp$, and $J_n^2(k_\perp v_\perp / \omega_c)$ for a representative loss-cone distribution.

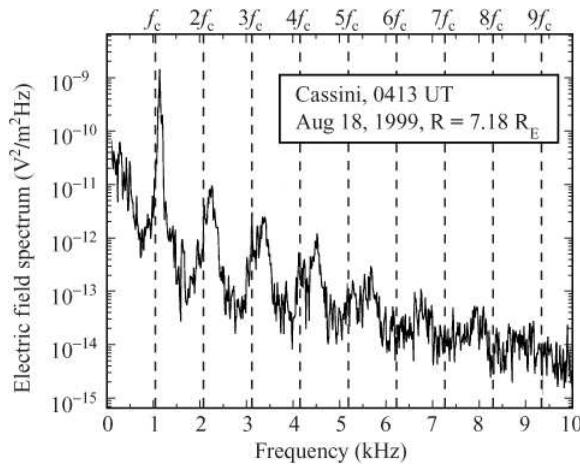


Fig. 9.16 A frequency–time spectrogram showing a series of electrostatic electron cyclotron emissions observed in the outer regions of Earth's magnetosphere.

Ion effects are, of course, already included in the general form of the Harris dispersion relation (9.2.20). Unfortunately, in the analysis of electrostatic ion cyclotron waves, it is not possible to omit the electron contribution in the same way that the ion contribution can be omitted in the analysis of electron cyclotron waves. However, the electron contribution to the dispersion relation can be greatly simplified by making use of suitable approximations. Since Landau damping is expected to strongly damp waves with parallel phase velocities near the electron thermal speed, we can assume that any unstable waves that might occur must have parallel phase velocities much greater than the electron thermal velocity. Therefore, we can use the high-phase velocity approximation, $k_{\parallel}v_{\parallel}/\omega \ll 1$, for the electron terms. Also, for comparable electron and ion temperatures, the electron cyclotron radius is much less than the ion cyclotron radius. Since the ion cyclotron waves are expected to have wavelengths comparable to the ion cyclotron radius, we can also use the long wavelength approximation ($k_\perp \rho_{ce} \ll 1$), for the electron Bessel function terms, which eliminates all of

these terms except for $n = 0$ and $n = \pm 1$. It is then easy to show that the electron contribution to the dispersion relation is identical to the low-temperature, long wavelength limit (9.2.22) derived earlier, i.e.,

$$D_e \approx D_{0e} = \left[1 - \frac{\omega_{pe}^2}{\omega^2} \right] \cos^2 \theta + \left[1 - \frac{\omega_{pe}^2}{\omega^2 - \omega_{ce}^2} \right] \sin^2 \theta. \quad (9.2.38)$$

If we assume that the ion cyclotron waves have frequencies near the low order harmonics of the ion cyclotron frequency, we also can make use of the approximations $\omega \ll \omega_{ce}$ and $\omega \ll \omega_{pe}$. Assuming that the plasma is also moderately dense ($\omega_{pe} \gg \omega_{ce}$), equation (9.2.38) simplifies to

$$D_e = \frac{\omega_{pe}^2}{\omega^2} \cos^2 \theta + \frac{\omega_{pe}^2}{\omega_{ce}^2} \sin^2 \theta. \quad (9.2.39)$$

To explore the conditions for instability, we again use the weak growth rate approximation. Following the usual procedure, the Plemelj relation is used to give the real and imaginary parts of the ion contribution to the dispersion relation, with the result that $D_{ion} = D_r + iD_i$. Since the ion plasma frequency is much smaller than the electron plasma frequency, the real part of the ion term is much smaller than the electron term and can be ignored. The real and imaginary parts of the dispersion relation are then given by

$$D_r = -\frac{\omega_{pe}^2}{\omega^2} \cos^2 \theta + \frac{\omega_{pe}^2}{\omega_{ce}^2} \sin^2 \theta, \quad (9.2.40)$$

and

$$D_i = -\pi \frac{k_{\parallel}}{|k_{\parallel}|} \frac{\omega_{pi}^2}{k^2} \sum_n \int_0^{\infty} J_n^2 \left(\frac{k_{\perp} v_{\perp}}{\omega_{ci}} \right) \times \left[k_{\parallel} \frac{\partial F_{i0}}{\partial v_{\parallel}} + \frac{n \omega_{ci}}{v_{\perp}} \frac{\partial F_{i0}}{\partial v_{\perp}} \right] 2\pi v_{\perp} dv_{\perp} \Big|_{v_{\parallel Res} = \frac{\omega - n \omega_{ci}}{k_{\parallel}}}. \quad (9.2.41)$$

As usual, the growth rate is given by

$$\gamma = \frac{-D_i}{\partial D_i / \partial \omega} \quad (9.2.42)$$

where ω and k are determined by $D_r(k, \omega) = 0$.

Since $\partial D_r / \partial \omega$ is easily shown to be positive from equation (9.2.40), the growth rate is positive whenever the integral in D_i is positive (again considering only $k_{\parallel} > 0$). As with the electron cyclotron instabilities, the possible free energy sources can be classified according to whether $\partial F_{i0} / \partial v_{\parallel}$ or $\partial F_{i0} / \partial v_{\perp}$ is positive. The prototypical free energy sources are then a “beam” and a “loss-cone,” just as in the case of the electron cyclotron instabilities. Of these, the loss-cone free energy source ($\partial F_{i0} / \partial v_{\perp} > 0$) is of special importance for both laboratory magnetic mirror machines and planetary radiation belts, since they always have

a loss cone. The region of the ion distribution function that causes the instability is essentially the same as in Figure 9.14, except that the parallel resonance velocity is now given by $v_{\parallel\text{Res}} = (\omega - n\omega_{ci})/k_{\parallel}$, and the resonance interaction is with the ions.

Since ω is expected to be much less than ω_{ce} , the condition $D_r = 0$, where D_r is given by equation (9.2.40), shows that the wave normal angle, θ , given by $\tan \theta = \omega_{ce}/\omega$, must be very close to $\pi/2$. The parallel wave number k_{\parallel} must then be quite small, which is consistent with the high-phase velocity approximation used for the electrons. In order for $v_{\parallel\text{Res}} = (\omega - n\omega_{ci})/k_{\parallel}$ to be in the range of the ion velocities, which are much smaller than the electron thermal velocities, it also follows that $\omega - n\omega_{ci}$ must be very small. The frequencies of the unstable modes are then very close to the harmonics of the ion cyclotron frequency, $\omega \approx n\omega_{ci}$. The wave normal angle is given approximately by

$$\tan \theta = \frac{\omega_{ce}}{\omega} \approx \frac{\omega_{ce}}{n\omega_{ci}} = \frac{m_i}{nm_e}. \quad (9.2.43)$$

For an electron-proton plasma, the wave normal angle for the $n = 1$ mode is approximately 89.97° , which is very nearly perpendicular to the magnetic field.

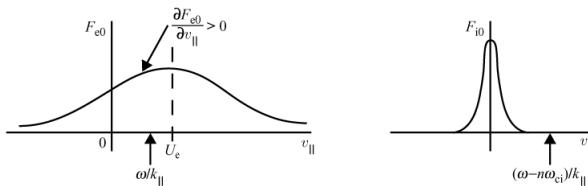


Fig. 9.17

Representative reduced one-dimensional electron and ion velocity distribution functions, F_{e0} and F_{i0} , used to illustrate the current-driven electrostatic ion cyclotron instability.

The range of perpendicular wave numbers that are unstable is again determined by the first maximum in $J_n^2(k_{\perp}v_{\perp}/\omega_{ci})$, which must coincide with the peak in $\partial F_{i0}/\partial v_{\perp}$ in order to give a net positive contribution to the v_{\perp} integral, very similar to the situation for the electrostatic electron cyclotron instabilities (see Figure 9.15).

9.2.4.2 Current-driven instabilities

Electrostatic ion cyclotron waves can also be driven unstable by currents flowing along the static magnetic field. Such magnetic field aligned currents are a common feature of magnetized plasmas and require that the electrons drift relative to the ions, as shown in Figure 9.17. The free energy source for the current-driven electrostatic ion cyclotron instability is the region of positive $\partial F_{e0}/\partial v_{\parallel}$ in the electron distribution, very similar to the current-driven ion acoustic instability (see Section 8.5.4.3). The existence of a current-driven ion cyclotron instability was first demonstrated by Drummond and Rosenbluth [1962], using Maxwellian electron and ion distribution functions of the form

$$F_{s0} = \left(\frac{m_s}{2\pi\kappa T_s} \right)^{3/2} \exp \left\{ -\frac{m_s[(v_{\parallel} - U_s)^2 + v_{\perp}^2]}{2\kappa T_s} \right\}, \quad (9.2.44)$$

where U_s is the drift velocity along the magnetic field. For a plasma consisting of electrons with a drift velocity U_e and one species of positive ion with a drift velocity $U_i = 0$, it is easy to show, using equations (??) and (9.2.27) that the Harris dispersion relation (9.2.20) can be written

$$\begin{aligned} D(k, p) = 1 & - \sum_n \frac{\Gamma_n(\beta_{ce})}{(k\lambda_{De})^2} \left[\frac{1}{2} Z'(\zeta_e - \zeta_{U_e}) - \frac{n\omega_{ce}}{k_{\parallel}} \sqrt{\frac{m_e}{2\kappa T_e}} Z(\zeta_e - \zeta_{U_e}) \right] \\ & - \sum_n \frac{\Gamma_n(\beta_{ci})}{(k\lambda_{Di})^2} \left[\frac{1}{2} Z'(\zeta_i) - \frac{n\omega_{ci}}{k_{\parallel}} \sqrt{\frac{m_i}{2\kappa T_i}} Z(\zeta_i) \right] = 0, \end{aligned} \quad (9.2.45)$$

where $Z(\zeta)$ is the plasma dispersion function described in Section ??, $\Gamma_n(\beta_{cs})$ is defined by equation (9.2.30), and for notational convenience, we introduce the following definitions

$$\zeta_s = \sqrt{\frac{m_s}{2\kappa T_s}} \frac{\omega - n\omega_{cs} + i\gamma}{k_{\parallel}} \quad \text{and} \quad \zeta_{U_e} = \sqrt{\frac{m_e}{2\kappa T_e}} U_e. \quad (9.2.46)$$

Since the wavelength of the electrostatic ion cyclotron waves is expected to be on the order of the ion cyclotron radius, which is usually much larger than the electron cyclotron radius, we can assume that $\beta_{ce} \ll 1$, so that $\Gamma_0(\beta_{ce}) = 1$ and $\Gamma_n(\beta_{ce}) = 0$ for $n \neq 0$. Also, since the ion cyclotron radius is usually much greater than the Debye length, we also assume that $k\lambda_{De} \ll 1$. The dispersion relation (9.2.45) then simplifies to

$$\frac{1}{2} Z'(\zeta_e - \zeta_{U_e}) + \frac{T_e}{T_i} \sum_n \Gamma_n(\beta_{ci}) \left[\frac{1}{2} Z'(\zeta_i) - \frac{n\omega_{ci}}{k_{\parallel}} \sqrt{\frac{m_i}{2\kappa T_i}} Z(\zeta_i) \right] = 0, \quad (9.2.47)$$

where we have made use of the relation

$$\frac{1}{(k\lambda_{Di})^2} = \frac{T_e}{T_i} \frac{1}{(k\lambda_{De})^2}. \quad (9.2.48)$$

To proceed further, we assume that the unstable k_{\parallel} values are such that the Landau resonance velocity for the electrons, $v_{\parallel Res} = \omega/k_{\parallel}$, is in the region of positive slope slightly below U_e (thereby providing a free energy source), and that the ion cyclotron resonance velocity, $v_{\parallel Res} = (\omega - \omega_{ci})/k_{\parallel}$, is well above the ion thermal speed (thereby avoiding strong ion cyclotron damping). These assumptions correspond to assuming that $|\zeta_e - \zeta_{U_e}| \ll 1$ and $|\zeta_i| \gg 1$. We can then use the small argument expansion for the plasma dispersion function in the electron term

$$Z'(\zeta_e - \zeta_{U_e}) \simeq -i \frac{k_{\parallel}}{|k_{\parallel}|} \sqrt{\pi} 2(\zeta_e - \zeta_{U_e}) - 2, \quad (9.2.49)$$

and the large-argument expansion in the ion term,

$$Z(\zeta_i) = i \frac{k_{\parallel}}{|k_{\parallel}|} \sqrt{\pi} e^{-\zeta_i^2} - \frac{1}{\zeta_i}, \quad \text{and} \quad Z'(\zeta_i) = -i \frac{k_{\parallel}}{|k_{\parallel}|} \sqrt{\pi} 2\zeta_i e^{-\zeta_i^2} + \frac{1}{\zeta_i^2}. \quad (9.2.50)$$

Using these expansions, equation (9.2.47) becomes

$$\begin{aligned} \frac{1}{2} \left[-i \frac{k_{\parallel}}{|k_{\parallel}|} \sqrt{\pi} 2(\zeta_e - \zeta_{U_e}) - 2 \right] + \frac{T_e}{T_i} \sum_n \Gamma_n(\beta_{ci}) \frac{1}{2} \left(-i \frac{k_{\parallel}}{|k_{\parallel}|} \sqrt{\pi} 2\zeta_i e^{-\zeta_i^2} \right) \\ - \frac{T_e}{T_i} \sum_n \Gamma_n(\beta_{ci}) \frac{n\omega_{ci}}{k_{\parallel}} \sqrt{\frac{m_i}{2\kappa T_i}} \left(i \frac{k_{\parallel}}{|k_{\parallel}|} \sqrt{\pi} e^{-\zeta_i^2} - \frac{1}{\zeta_i} \right) = 0. \end{aligned} \quad (9.2.51)$$

Writing $\zeta_i = x_i + iy_i$ where

$$x_i = \sqrt{\frac{m_i}{2\kappa T_i}} \frac{\omega - n\omega_{ci}}{k_{\parallel}} \quad \text{and} \quad y_i = \sqrt{\frac{m_i}{2\kappa T_i}} \frac{\gamma}{k_{\parallel}} \quad (9.2.52)$$

and assuming that the growth rate is small (i.e., $|y_i| \ll |x_i|$), it is easy to show that the real and imaginary parts of the dispersion relation (9.2.51) yield two equations. The real part gives an equation for the wave frequencies

$$\frac{\omega - n\omega_{ci}}{n\omega_{ci}} = \frac{T_e}{T_i} \Gamma_n(\beta_{ci}), \quad (9.2.53)$$

and the imaginary part (after eliminating the x_i term using the real part of the dispersion relation) gives the growth rate

$$\begin{aligned} \frac{\gamma}{n\omega_{ci}} &= \frac{k_{\parallel}}{|k_{\parallel}|} \Gamma_n(\beta_{ci}) \frac{T_e}{T_i} \sqrt{\frac{\pi}{2}} \\ &\times \left[\frac{U_e}{C_e} - \left\{ \frac{1}{C_e} + \frac{T_e}{T_i} \frac{\Gamma_n(\beta_{ci})}{C_i} \exp \left[-\frac{1}{2} \left(\frac{\omega - n\omega_{ci}}{k_{\parallel} C_i} \right)^2 \right] \right\} \frac{\omega}{k_{\parallel}} \right], \end{aligned} \quad (9.2.54)$$

where C_e and C_i are the electron and ion thermal speeds, as defined by equation (??). Note that the growth rate is proportional to $\Gamma_n(\beta_{ci})$, which has the form shown in Figure 9.6.

Since Γ_1 has the largest peak value (0.22 at $\beta_{ci} = 1.24$), the $n = 1$ mode has the highest growth rate. Also note that since the peak values of Γ_n decrease rapidly with increasing n and are always less than 0.22, the frequencies of all the modes are restricted to a narrow range of frequencies slightly above the corresponding harmonic, $n\omega_{ci}$, of the ion cyclotron frequency.

The critical drift velocity, U_e^* , for the onset ($\gamma = 0$) of the current-driven electrostatic ion cyclotron instability can be obtained by setting the term in the outermost bracket in equation (9.2.54) to zero, which gives

$$\frac{U_e^*}{C_e} = \frac{\omega}{k_{\parallel} C_e} + \frac{T_e}{T_i} \frac{0.22\omega}{k_{\parallel} C_i} \exp \left[-\frac{1}{2} \left(\frac{\omega - n\omega_{ci}}{k_{\parallel} C_i} \right)^2 \right], \quad (9.2.55)$$

where we have used $\Gamma_1(1.24) = 0.22$ for $\Gamma_n(\beta_{ci})$, since the $n = 1$ mode has the highest

growth rate and is therefore the most unstable. The critical drift velocity varies with k_{\parallel} . For small k_{\parallel} (i.e., ω/k_{\parallel} near but less than U_e), the threshold is very high, corresponding to the fact that the slope of the electron velocity distribution, which provides the free energy source for the instability, is small in this region (see Figure 9.17). As k_{\parallel} increases and the slope of the electron velocity distribution function at $v_{\parallel\text{Res}} = \omega/k_{\parallel}$ increases, the threshold drift velocity decreases. However, as k_{\parallel} continues to increase, the ion cyclotron resonance velocity, $v_{\parallel\text{Res}} = (\omega - n\omega_{ci})/k_{\parallel}$, begins to move downward into the main part of the ion distribution function (see Figure 9.17), causing an increase in ion cyclotron damping, which eventually raises the threshold drift velocity. These variations lead to a broad minimum in the threshold drift velocity. Although this minimum is difficult to analyze analytically, it is easily evaluated using numerical techniques. Drummond and Rosenbluth [1962] have shown that the threshold drift velocity is given approximately by

$$U_e^* = [14(T_i/T_e) + 3]C_i. \quad (9.2.56)$$

For moderate electron-to-ion temperature ratios ($T_e/T_i \sim 1$), the threshold electron drift velocity for the current-driven electrostatic ion cyclotron instability is very low, much lower than the corresponding threshold for the current-driven ion acoustic instability given by equation (??). Because of the low threshold for this instability, electrostatic ion cyclotron waves are often observed in both laboratory and space plasmas in regions where field aligned currents are present. An example of a series of current-driven electrostatic ion cyclotron waves detected in a laboratory Q machine is shown in Figure 9.18.

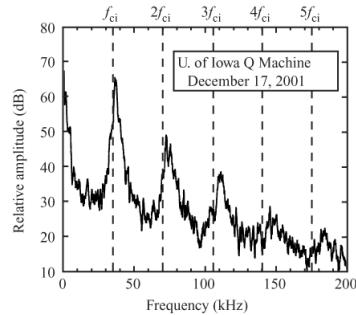


Fig. 9.18 A spectrum of current-driven electrostatic ion cyclotron waves observed in a laboratory Q machine, using a cesium plasma [plot provided by R. Merlino].

9.3 Electromagnetic waves

Since an electromagnetic wave has both an electric field and a magnetic field, it is not possible to describe the electric field by an electrostatic potential, as was done in the previous section. Instead, the full set of Maxwell's equations must be used. The basic procedure used

to analyze electromagnetic wave propagation in a dielectric medium has been described in Chapter ??, and is based on finding non-trivial solutions to the homogeneous equation

$$\mathbf{k} \times (\mathbf{k} \times \tilde{\mathbf{E}}) + \frac{\omega^2}{c^2} \overset{\leftrightarrow}{\mathbf{K}} \cdot \tilde{\mathbf{E}} = 0, \quad (9.3.1)$$

where $\tilde{\mathbf{E}}$ is the electric field of the wave and $\overset{\leftrightarrow}{\mathbf{K}}$ is the dielectric tensor. Non-trivial solutions of this system of linear equations are possible only when the determinant of the matrix in the homogeneous equation (9.3.1) is zero, which gives the dispersion relation. Before we can proceed further, we must compute the dielectric tensor. The procedure used to compute the dielectric tensor consists of first solving the linearized Vlasov equation for the first-order distribution function, \tilde{f}_s , for some given electric field, $\tilde{\mathbf{E}}$, and then using \tilde{f}_s to compute the current density

$$\tilde{\mathbf{J}} = \sum_s e_s \int_{-\infty}^{\infty} \mathbf{v} \tilde{f}_s d^3 v. \quad (9.3.2)$$

By expressing the current density in the form $\tilde{\mathbf{J}} = \overset{\leftrightarrow}{\sigma} \cdot \tilde{\mathbf{E}}$ the conductivity tensor $\overset{\leftrightarrow}{\sigma}$ can be identified. Once the conductivity tensor is known, the dielectric tensor can be computed using $\overset{\leftrightarrow}{\mathbf{K}} = \mathbf{1} - \overset{\leftrightarrow}{\sigma}/(i\omega\epsilon_0)$, see equation (??).

9.3.1 The dispersion relation

Following the above procedure, the first step in deriving the dispersion relation is to solve the linearized Vlasov equation (9.1.13) for the first-order distribution, \tilde{f}_s , for an arbitrary wave electric field, $\tilde{\mathbf{E}}$. The procedure is basically the same as for the electrostatic waves described in the previous section. However, the solution is more complicated because of the $\mathbf{v} \times (\mathbf{k} \times \tilde{\mathbf{E}})$ term introduced by the magnetic wave field. This term can be expanded using the BAC minus CAB rule, which gives

$$\mathbf{v} \times (\mathbf{k} \times \tilde{\mathbf{E}}) = \mathbf{k}(\mathbf{v} \cdot \tilde{\mathbf{E}}) - (\mathbf{v} \cdot \mathbf{k})\tilde{\mathbf{E}}. \quad (9.3.3)$$

Substituting the above term into equation (9.1.13) gives the following linear differential equation

$$\begin{aligned} \frac{\partial \tilde{f}_s}{\partial \phi} - i(\alpha_s + \beta_s \cos \phi)\tilde{f}_s \\ = \frac{e_s}{m_s \omega_{cs}} \left[\left(1 - \frac{k_{||}v_{||}}{\omega} - \frac{k_{\perp}v_{\perp}}{\omega} \cos \phi \right) \tilde{\mathbf{E}} + \frac{\mathbf{k}}{\omega} (\mathbf{v} \cdot \tilde{\mathbf{E}}) \right] \cdot \nabla_{\mathbf{v}} f_{s0}. \end{aligned} \quad (9.3.4)$$

As before, to solve this equation for \tilde{f}_s we must express all of the terms on the right-hand side in cylindrical velocity space coordinates (v_{\perp} , and $\phi, v_{||}$). Starting with rectangular (v_x, v_y, v_z) coordinates, the dot-product between $\tilde{\mathbf{E}}$ and $\nabla_{\mathbf{v}} f_{s0}$ in the first term in the brackets can be written

$$\begin{aligned}\tilde{\mathbf{E}} \cdot \nabla_{\mathbf{v}} f_{s0} &= \tilde{E}_x \frac{\partial f_{s0}}{\partial v_x} + \tilde{E}_y \frac{\partial f_{s0}}{\partial v_y} + \tilde{E}_z \frac{\partial f_{s0}}{\partial v_z} \\ &= \tilde{E}_x \frac{\partial v_{\perp}}{\partial v_x} \frac{\partial f_{s0}}{\partial v_{\perp}} + \tilde{E}_y \frac{\partial v_{\perp}}{\partial v_y} \frac{\partial f_{s0}}{\partial v_{\perp}} + \tilde{E}_z \frac{\partial f_{s0}}{\partial v_{\parallel}}.\end{aligned}\quad (9.3.5)$$

From the transformation equation $v_{\perp} = v_x \cos \phi + v_y \sin \phi$, it is easy to show that

$$\frac{\partial v_{\perp}}{\partial v_x} = \frac{v_x}{v_{\perp}} = \cos \phi \quad \text{and} \quad \frac{\partial v_{\perp}}{\partial v_y} = \frac{v_y}{v_{\perp}} = \sin \phi. \quad (9.3.6)$$

Substituting these relations into equation (9.3.5) then gives

$$\tilde{\mathbf{E}} \cdot \nabla_{\mathbf{v}} f_{s0} = \frac{\partial f_{s0}}{\partial v_{\perp}} (\tilde{E}_x \cos \phi + \tilde{E}_y \sin \phi) + \frac{\partial f_{s0}}{\partial v_{\parallel}} \tilde{E}_z. \quad (9.3.7)$$

Similarly, the dot-product between \mathbf{v} and $\tilde{\mathbf{E}}$ in the second term in the brackets of equation (9.3.4) can be written

$$\begin{aligned}\mathbf{v} \cdot \tilde{\mathbf{E}} &= v_x \tilde{E}_x + v_y \tilde{E}_y + v_z \tilde{E}_z \\ &= v_{\perp} (\tilde{E}_x \cos \phi + \tilde{E}_y \sin \phi) + v_{\parallel} \tilde{E}_z.\end{aligned}\quad (9.3.8)$$

Substituting equations (9.3.7) and (9.3.8) into equation (9.3.4), and changing the zero-order distribution function to the normalized zero-order distribution function using $f_{s0} = n_{s0} F_{s0}$ gives the equation

$$\begin{aligned}\frac{\partial \tilde{f}_s}{\partial \phi} - i(\alpha_s + \beta_s \cos \phi) \tilde{f}_s \\ = \frac{e_s n_s}{m_s \omega_{cs}} \left[(A_s \tilde{E}_x + D_s \tilde{E}_z) \cos \phi + A_s \tilde{E}_y \sin \phi + \frac{\partial F_{s0}}{\partial v_{\parallel}} \tilde{E}_z \right],\end{aligned}\quad (9.3.9)$$

where to simplify the notation we have introduced the definitions

$$A_s = \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}}{\omega} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right) \quad (9.3.10)$$

and

$$D_s = -\frac{k_{\perp}}{\omega} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right). \quad (9.3.11)$$

This linear differential equation can be solved using the same integrating factor used to solve equation (9.2.1). The result is

$$\begin{aligned}\tilde{f}_s &= i \frac{e_s n_s}{m_s \omega_{cs}} \sum_{m,n} \frac{e^{i(m-n)\phi}}{\alpha_s + n} J_m(\beta_s) \\ &\times \left[\left(\frac{n}{\beta_s} J_n(\beta_s) \tilde{E}_x + i J'_n(\beta_s) \tilde{E}_y \right) A_s + J_n(\beta_s) B_s \tilde{E}_z \right],\end{aligned}\quad (9.3.12)$$

where the Bessel function identities

$$J_{n+1}(\beta_s) + J_{n-1}(\beta_s) = \frac{2n}{\beta_s} J_n(\beta_s) \quad \text{and} \quad J_{n+1}(\beta_s) - J_{n-1}(\beta_s) = -2 J'_n(\beta_s) \quad (9.3.13)$$

have been used, and the term B_s is defined by

$$B_s = \frac{\partial F_{s0}}{\partial v_{\parallel}} + \frac{n}{\beta_s} D_s = \frac{\partial F_{s0}}{\partial v_{\parallel}} - \frac{n \omega_{cs}}{\omega v_{\perp}} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right). \quad (9.3.14)$$

Next, we compute the current by evaluating the integral

$$\tilde{\mathbf{J}} = \sum_s e_s \int_0^{2\pi} \int_{-\infty}^{\infty} \int_0^{\infty} \mathbf{v} \tilde{f}_s v_{\perp} dv_{\perp} dv_{\parallel} d\phi, \quad (9.3.15)$$

where $\mathbf{v} = \hat{\mathbf{x}} v_{\perp} \cos \phi + \hat{\mathbf{y}} v_{\perp} \sin \phi + \hat{\mathbf{z}} v_{\parallel}$. The result after integrating over $d\phi$ can be organized in matrix form as

$$\begin{bmatrix} \tilde{J}_x \\ \tilde{J}_y \\ \tilde{J}_z \end{bmatrix} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix}, \quad (9.3.16)$$

where the 3×3 matrix is the conductivity tensor, $\vec{\sigma}$. The integral in (9.3.15) is most easily performed by expressing $\cos \phi$ and $\sin \phi$ in complex exponential form and noting that

$$\int_0^{2\pi} e^{i(m-n)\phi} d\phi = 2\pi \delta_{m,n} \quad \text{and} \quad \int_0^{2\pi} e^{i(m-n\pm 1)\phi} d\phi = 2\pi \delta_{m,n\pm 1} \quad (9.3.17)$$

where $\delta_{m,n}$ is the Kronecker delta.

The Kronecker deltas have the effect of collapsing the double sum over m and n into a single sum over n , very similar to the procedure used to derive the Harris dispersion relation (9.2.20). The conductivity tensor is then found to be

$$\begin{aligned}\vec{\sigma} &= i \sum_s \frac{e_s^2 n_s}{m_s \omega_{cs}} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{2\pi v_{\perp} dv_{\perp} dv_{\parallel}}{\alpha_s + n} \\ &\times \begin{bmatrix} A_s \frac{n^2 v_{\perp}}{\beta_s^2} J_n^2 & i A_s \frac{nv_{\perp}}{\beta_s} J_n J'_n & B_s \frac{nv_{\perp}}{\beta_s} J_n^2 \\ -i A_s \frac{nv_{\perp}}{\beta_s} J_n J'_n & A_s v_{\perp} J'_n J'_n & -i B_s v_{\perp} J_n J'_n \\ A_s \frac{nv_{\parallel}}{\beta_s} J_n^2 & i A_s v_{\parallel} J_n J'_n & B_s v_{\parallel} J_n^2 \end{bmatrix},\end{aligned}\quad (9.3.18)$$

where the Bessel function identities given in (9.3.13) have been used.

Having computed the conductivity tensor, it is then an easy matter to compute the dielectric tensor, which is given by $\vec{\mathbf{K}} = \vec{\mathbf{1}} - \vec{\sigma}/(i\omega\epsilon_0)$. After substituting the appropriate expressions for α_s, β_s, A_s , and B_s from equations (9.1.14), (9.3.10), and (9.3.14) into equation (9.3.18), the dielectric tensor elements are found to be

$$K_{xx} = 1 - \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{n^2 J_n^2(\beta_s)}{\beta_s^2(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\left(1 - \frac{k_{\parallel}v_{\parallel}}{\omega} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}v_{\perp}}{\omega} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] 2\pi v_{\perp}^2 dv_{\perp} dv_{\parallel}, \quad (9.3.19)$$

$$K_{xy} = -i \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{n J_n(\beta_s) J'_n(\beta_s)}{\beta_s(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\left(1 - \frac{k_{\parallel}v_{\parallel}}{\omega} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}v_{\perp}}{\omega} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] 2\pi v_{\perp}^2 dv_{\perp} dv_{\parallel}, \quad (9.3.20)$$

$$K_{xz} = - \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{n J_n^2(\beta_s)}{\beta_s(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\frac{\partial F_{s0}}{\partial v_{\parallel}} - \frac{n\omega_{cs}}{\omega v_{\perp}} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right) \right] 2\pi v_{\perp}^2 dv_{\perp} dv_{\parallel}, \quad (9.3.21)$$

$$K_{yx} = -i \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{n J_n(\beta_s) J'_n(\beta_s)}{\beta_s(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\left(1 - \frac{k_{\parallel}v_{\parallel}}{\omega} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}v_{\perp}}{\omega} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] 2\pi v_{\perp}^2 dv_{\perp} dv_{\parallel}, \quad (9.3.22)$$

$$K_{yy} = 1 - \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{J'_n(\beta_s) J'_n(\beta_s)}{(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\left(1 - \frac{k_{\parallel}v_{\parallel}}{\omega} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}v_{\perp}}{\omega} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] 2\pi v_{\perp}^2 dv_{\perp} dv_{\parallel}, \quad (9.3.23)$$

$$K_{yz} = i \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{J_n(\beta_s) J'_n(\beta_s)}{(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\frac{\partial F_{s0}}{\partial v_{\parallel}} - \frac{n\omega_{cs}}{\omega v_{\perp}} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right) \right] 2\pi v_{\perp}^2 dv_{\perp} dv_{\parallel}, \quad (9.3.24)$$

$$K_{zx} = - \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{n J_n^2(\beta_s)}{\beta_s(k_{\parallel}v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\left(1 - \frac{k_{\parallel}v_{\parallel}}{\omega} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}v_{\perp}}{\omega} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] 2\pi v_{\perp} v_{\parallel} dv_{\perp} dv_{\parallel}, \quad (9.3.25)$$

$$K_{zy} = -i \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{J_n(\beta_s) J'_n(\beta_s)}{(k_{\parallel} v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\left(1 - \frac{k_{\parallel} v_{\parallel}}{\omega} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel} v_{\perp}}{\omega} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] 2\pi v_{\perp} v_{\parallel} dv_{\perp} dv_{\parallel}, \quad (9.3.26)$$

and

$$K_{zz} = 1 - \sum_s \frac{\omega_{ps}^2}{\omega} \sum_n \int_{-\infty}^{\infty} \int_0^{\infty} \frac{J_n^2(\beta_s)}{(k_{\parallel} v_{\parallel} - \omega + n\omega_{cs})} \\ \times \left[\frac{\partial F_{s0}}{\partial v_{\parallel}} - \frac{n\omega_{cs}}{\omega v_{\perp}} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right) \right] 2\pi v_{\perp} v_{\parallel} dv_{\perp} dv_{\parallel}, \quad (9.3.27)$$

The homogeneous equation for the electric field is obtained by substituting the dielectric tensor elements into equation (9.3.1), which becomes

$$\begin{bmatrix} K_{xx} - \frac{c^2 k^2}{\omega^2} \cos^2 \theta & K_{xy} & k_{xz} + \frac{c^2 k^2}{\omega^2} \sin \theta \cos \theta \\ K_{yx} & K_{yy} - \frac{c^2 k^2}{\omega^2} & K_{yz} \\ K_{zx} + \frac{c^2 k^2}{\omega^2} \sin \theta \cos \theta & K_{zy} & K_{zz} - \frac{c^2 k^2}{\omega^2} \sin \theta \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (9.3.28)$$

The dispersion relation is then given by the determinant of the matrix. Although the determinant can, in principle, be evaluated, the resulting equation is too complicated to be usefully displayed, since each term consists of an infinite series of integrals involving Bessel functions. To evaluate the dispersion relation it is usually necessary to use either numerical methods or highly specialized assumptions. However, one special case that can be analyzed analytically is when the wave vector is parallel to the magnetic field ($\theta = 0$). Since parallel propagation illustrates many of the important features that occur for arbitrary wave normal angles, we will analyze this case in some detail.

9.3.2 Parallel propagation

From inspection of the homogeneous equation (9.3.28), it is evident that the dispersion relation simplifies considerably if the wave vector is parallel to the magnetic field (i.e., $\theta = 0$). There are three such simplifications. First, it is easy to see that the terms involving $\sin \theta$ are all zero. Second, since $\beta_s = k_{\perp} v_{\perp} / \omega_{cs} = 0$ for parallel propagation, it is also easy to see that the K_{xz} and K_{zx} terms are zero, since all of the Bessel function terms, $n J_n^2(\beta_s)$, go to zero as $\beta_s \rightarrow 0$. Third, it turns out that the K_{yz} and K_{zy} terms are also zero. That these terms are zero can be seen by noting that, as β_s goes to zero, all of the $J_n(\beta_s) J'_n(\beta_s)$ terms go to zero (note that $J_0(\beta_s) \simeq 1$ and $J'_0(\beta_s) = -J_1(\beta_s) \simeq -\beta_s/2$, as $\beta_s \rightarrow 0$). The homogeneous equation (9.3.28) then simplifies to

$$\begin{bmatrix} K_{xx} - \frac{c^2 k^2}{\omega^2} & K_{xy} & 0 \\ K_{yx} & K_{yy} - \frac{c^2 k^2}{\omega^2} & 0 \\ 0 & 0 & K_{zz} \end{bmatrix} \begin{bmatrix} \tilde{E}_x \\ \tilde{E}_y \\ \tilde{E}_z \end{bmatrix} = 0. \quad (9.3.29)$$

By taking the determinant of the matrix in the above equation, the dispersion relation is easily shown to be

$$\mathcal{D}(k, \omega) = \left[\left(K_{xx} - \frac{c^2 k^2}{\omega^2} \right) \left(K_{yy} - \frac{c^2 k^2}{\omega^2} \right) - K_{xy} K_{yx} \right] K_{zz} = 0. \quad (9.3.30)$$

From the homogeneous equation (9.3.29), it can be shown that the roots of the dispersion relation separate into three types: a series of electrostatic modes with $\hat{\mathbf{E}} = (0, 0, \hat{E}_z)$; a series of electromagnetic modes with eigenvectors $\hat{\mathbf{E}} = (\hat{E}_0, i\hat{E}_0, 0)$; and a series of electromagnetic modes with eigenvectors $\hat{\mathbf{E}} = (\hat{E}_0, -i\hat{E}_0, 0)$. The electrostatic modes are the same as those described in Chapter 8 for an unmagnetized plasma. For these modes, the zero-order magnetic field plays no role, since the particle motions are along the magnetic field and there is no $\mathbf{v} \times \hat{\mathbf{B}}_0$ force. The analysis of the remaining electromagnetic modes can be simplified considerably by noting that $K_{xy} = -K_{yx}$ and $K_{xx} = K_{yy}$. The latter occurs because, in the limit $\beta_s \rightarrow 0$, the $n^2 J_n^2(\beta_s)/\beta_s^2$ terms in K_{xx} and the $J'_n(\beta_s)J'_n(\beta_s)$ terms in K_{yy} both reduce to 1/4 for $n = \pm 1$, and zero for $n = 0$. With these simplifications, the electromagnetic part of the dispersion reduces to

$$\mathcal{D}(k, \omega) = K_{xx} - \frac{c^2 k_{\parallel}^2}{\omega^2} \pm iK_{xy} = 0 \quad (9.3.31)$$

which, after substituting the appropriately simplified expressions for K_{xx} and K_{xy} , can be written in the form

$$\begin{aligned} \mathcal{D}(k, \omega) &= 1 - \frac{c^2 k_{\parallel}^2}{\omega^2} - \sum_s \frac{\omega_{ps}^2}{\omega} \\ &\times \int_{-\infty}^{\infty} \int_0^{\infty} \frac{\frac{\partial F_{s0}}{\partial v_{\perp}} + \frac{k_{\parallel}}{\omega} \left(v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} - v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} \right)}{k_{\parallel} v_{\parallel} - \omega \pm \omega_{cs}} \pi v_{\perp}^2 dv_{\perp} dv_{\parallel} = 0. \end{aligned} \quad (9.3.32)$$

After factoring out k_{\parallel}/ω from both the numerator and the denominator, the above equation can be written in the following somewhat simpler form

$$\mathcal{D}(k, \omega) = 1 - \frac{c^2 k_{\parallel}^2}{\omega^2} - \sum_s \frac{\omega_{ps}^2}{\omega} \int_{-\infty}^{\infty} \frac{G_{s0}(v_{\parallel})}{v_{\parallel} - \frac{\omega \pm \omega_{cs}}{k_{\parallel}}} dv_{\parallel} = 0. \quad (9.3.33)$$

where, for convenience, we have defined a reduced one-dimensional distribution function

$$G_{s0}(v_{\parallel}) = \int_0^{\infty} \left[\left(\frac{\omega}{k_{\parallel}} - v_{\parallel} \right) \frac{\partial F_{s0}}{\partial v_{\perp}} + v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] \pi v_{\perp}^2 dv_{\perp}. \quad (9.3.34)$$

Finally, the above equations can be converted to Laplace transform notation by changing ω to $i\rho$ and changing the v_{\parallel} integration contour such that it passes below the pole at $v_{\parallel} = (i\rho \pm \omega_{cs})/k_{\parallel}$ for $k_{\parallel} > 0$, and above the pole for $k_{\parallel} < 0$, as in Figure 9.13.

9.3.2.1 The low-temperature, long wavelength limit

As with the electrostatic dispersion relation, it is useful to consider the low-temperature, long-wavelength limit, so that we can relate the dispersion relation to the results from cold plasma theory in Chapter 4. This can be done by expanding the integrand in equation (9.3.33) in powers of $k_{\parallel}v_{\parallel}/(\omega \pm \omega_{\text{cs}})$ and considering the low-temperature, long wavelength limit of the resulting series of integrals. However, a simpler approach is to assume that the plasma is cold, with a distribution function given by

$$F_{s0} = \frac{1}{2\pi v_{\perp}} \delta(v_{\perp}) \delta(v_{\parallel}). \quad (9.3.35)$$

After substituting this distribution function into equation (9.3.34) and integrating by parts once, the reduced one-dimensional distribution function simplifies to

$$G_{s0}(v_{\parallel}) = -\left(\frac{\omega}{k_{\parallel}} - v_{\parallel}\right) \delta(v_{\parallel}). \quad (9.3.36)$$

Completing the v_{\parallel} integration in equation (9.3.33) then gives the dispersion relation

$$D(k, \omega) = 1 - \frac{c^2 k_{\parallel}^2}{\omega^2} - \sum_s \frac{\omega_{\text{ps}}^2}{\omega(\omega \pm \omega_{\text{cs}})} = 0. \quad (9.3.37)$$

Noting that $n^2 = c^2 k_{\parallel}^2 / \omega^2$, the above equation is seen to be identical to the cold plasma dispersion relation obtained in Chapter ?? for electromagnetic waves propagating parallel to the magnetic field ($\theta = 0$). The plus sign corresponds to right-hand polarized waves ($n^2 = R$), and the minus sign corresponds to left-hand polarized waves ($n^2 = L$).

9.3.2.2 Cyclotron damping

Next we show that the imaginary part of the dispersion relation introduced by the resonance at $v_{\parallel} = (\text{i}p \pm \omega_{\text{cs}})/k_{\parallel}$ causes electromagnetic waves to be damped, similar in some respects to Landau damping. This damping is called cyclotron damping. In the weak damping limit, the frequency and growth rate can be determined from equation (9.3.33) in the usual way using the Plemelj formula, which gives

$$D_r = 1 - \frac{c^2 k_{\parallel}^2}{\omega^2} - \sum_s \frac{\omega_{\text{ps}}^2}{\omega^2} P \int_{-\infty}^{\infty} \frac{G_{s0}(v_{\parallel})}{v_{\parallel} - \frac{\omega \pm \omega_{\text{cs}}}{k_{\parallel}}} dv_{\parallel} = 0 \quad (9.3.38)$$

and

$$\gamma = \frac{-D_r}{\partial D_r / \partial \omega} = \pi \frac{k_{\parallel} / |k_{\parallel}|}{\partial D_r / \partial \omega} \sum_s \frac{\omega_{\text{ps}}^2}{\omega^2} G_{s0}(v_{\parallel}) \Big|_{v_{\parallel}=v_{\parallel\text{Res}}}, \quad (9.3.39)$$

where $G_{s0}(v_{\parallel})$ is given by equation (9.3.34), and $v_{\parallel\text{Res}} = (\omega \pm \omega_{\text{cs}})/k_{\parallel}$ is the cyclotron resonance velocity. The term $k_{\parallel} / |k_{\parallel}|$ takes care of the usual sign change that occurs in the residue as k_{\parallel} changes sign.

Equation (9.3.39) shows that the growth or damping of the wave is determined by the sign of $\sum_s \omega_{ps}^2 G_{s0}(v_{\parallel Res})$. For the purpose of evaluating $G_{s0}(v_{\parallel Res})$, it is convenient to separate this function into two integrals as follows

$$G_{s0}(v_{\parallel}) = \frac{\omega}{k_{\parallel}} \int_0^{\infty} \frac{\partial F_{s0}}{\partial v_{\perp}} \pi v_{\perp}^2 dv_{\perp} - \int_0^{\infty} \left(v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} - v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right) \pi v_{\perp}^2 dv_{\perp}. \quad (9.3.40)$$

The first integral in the above equation can be integrated by parts once to give

$$G_{s0}(v_{\parallel}) = \frac{\omega}{k_{\parallel}} \int_0^{\infty} F_{s0} 2\pi v_{\perp} dv_{\perp} - \int_0^{\infty} \left(v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} - v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right) \pi v_{\perp}^2 dv_{\perp}. \quad (9.3.41)$$

which, when substituted into equation (9.3.39), gives the following equation for the growth rate

$$\begin{aligned} \gamma = & \pi \frac{1}{\partial D_r / \partial \omega} \sum_s \frac{\omega_{ps}^2}{\omega^2} \left[- \frac{\omega}{|k_{\parallel}|} \int_0^{\infty} F_{s0} 2\pi v_{\perp} dv_{\perp} \right. \\ & \left. - \frac{k_{\parallel}}{|k_{\parallel}|} \int_0^{\infty} \left(v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} - v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right) \pi v_{\perp}^2 dv_{\perp} \right] \Big|_{v_{\parallel}=v_{\parallel Res}}. \end{aligned} \quad (9.3.42)$$

To help understand the various terms in the above equation, we next consider the special case of an isotropic velocity distribution function. For an isotropic distribution, the distribution function is a function only of the magnitude of the velocity, $F_{s0} = F_{s0}(v)$, where $v = (v_{\parallel}^2 + v_{\perp}^2)^{1/2}$. For such a distribution it is easy to show that

$$\frac{\partial F_{s0}}{\partial v_{\perp}} = \frac{\partial v}{\partial v_{\perp}} \frac{\partial F_{s0}}{\partial v} = \frac{v_{\perp}}{v} \frac{\partial F_{s0}}{\partial v}, \quad \text{and} \quad \frac{\partial F_{s0}}{\partial v_{\parallel}} = \frac{\partial v}{\partial v_{\parallel}} \frac{\partial F_{s0}}{\partial v} = \frac{v_{\parallel}}{v} \frac{\partial F_{s0}}{\partial v}. \quad (9.3.43)$$

After substituting these expressions into equation (9.3.42), it is evident that the two terms in the second integral cancel. The growth rate is then given by

$$\gamma = -\pi \frac{\omega / |k_{\parallel}|}{\partial D_r / \partial \omega} \sum_s \frac{\omega_{ps}^2}{\omega^2} \int_0^{\infty} F_{s0} 2\pi v_{\perp} dv_{\perp} \Big|_{v_{\parallel}=v_{\parallel Res}}, \quad (9.3.44)$$

which is negative, indicating damping, since, as we will show, $\partial D_r / \partial \omega$ is always positive. This damping is called cyclotron damping. Although cyclotron damping has many similarities to Landau damping, it differs from Landau damping in one important respect. As can be seen from the above equation, cyclotron damping is directly proportional to the distribution function evaluated at the parallel resonance velocity, whereas Landau damping is proportional to the slope of the distribution function at the resonance velocity. This difference indicates that cyclotron damping is fundamentally different from Landau damping in that all particles in the vicinity of the resonance contribute to the damping, not just those below the resonance velocity as in the case of Landau damping.

If we limit our consideration to resonance velocities well above the thermal velocity, the real part of the dispersion relation, equation (9.3.38), reduces to the same equations

($n^2 = R$ and $n^2 = L$) obtained in a cold plasma theory, i.e., equation (9.3.37). Thus, if the cyclotron resonance velocity is not too close to the main part of the distribution function, the waves propagate exactly as in cold plasma theory, except for the damping, which is directly proportional to the number of particles in resonance with the wave. For example, consider the whistler mode, for which the ion motions can be ignored. If the electrons have a Maxwellian velocity distribution with temperature T , the damping ($-\gamma$) is given by

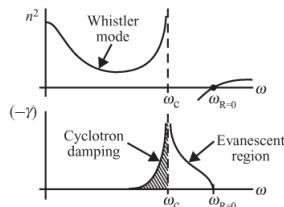


Fig. 9.19 The damping ($-\gamma$) of the whistler mode as a function of frequency. Cyclotron damping becomes very strong just below the electron cyclotron frequency (the hatched region). The whistler mode cannot propagate in the evanescent region from ω_c to $\omega_{R=0}$.

$$(-\gamma) = \pi \frac{\omega/|k_{||}|}{\partial D_r / \partial \omega} \frac{\omega_p^2}{\omega^2} \left(\frac{m}{2\pi\kappa T} \right)^{1/2} \exp \left(-\frac{mv_{||Res}^2}{2\kappa T} \right) \Bigg|_{v_{||Res} = \frac{\omega - \omega_c}{k_{||}}}. \quad (9.3.45)$$

A plot of the index of refraction squared, $n^2 = c^2 k^2 / \omega^2$, and the damping rate ($-\gamma$) for the right-hand polarized whistler mode, $n^2 = R$, is shown in Figure 9.19. Because the damping rate ($-\gamma$) has an exponential dependence on minus $(\omega - \omega_c)^2$, cyclotron damping is strong only near the electron cyclotron frequency. Note that the frequency range over which significant damping occurs decreases with decreasing temperature, and goes to zero for a cold plasma ($T = 0$). Also note that in the frequency range between ω_c and $\omega_{R=0}$, the imaginary part of the frequency ($-\gamma$) is non-zero even when the plasma is cold. This is the evanescent region discussed in Chapter ???. Therefore, the reasons for the non-zero values of ($-\gamma$) are qualitatively different above and below the cyclotron frequency. Below the cyclotron frequency, the damping ($-\gamma$) is large because the plasma absorbs energy from the wave via cyclotron damping, whereas in the region above the cyclotron frequency the wave cannot propagate because of dynamical considerations (the displacement current in Maxwell's equation has the wrong sign).

Similar considerations also apply to the left-hand polarized ion cyclotron mode. The damping is then called ion cyclotron damping. Ion cyclotron whistlers provide a good example of ion cyclotron damping. Ion cyclotron whistlers typically exhibit an abrupt cutoff in the wave amplitude as the wave frequency asymptotically approaches the ion cyclotron frequency (see Figure ??). The abrupt decrease in the wave amplitude for a proton whistler is shown in Figure 9.20 as the difference between the wave frequency and the proton cyclotron frequency, Δf , goes to zero.

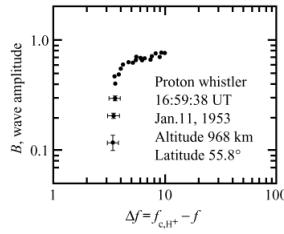


Fig. 9.20 An illustration showing the abrupt onset of ion cyclotron damping as the frequency of a proton whistler approaches the proton cyclotron frequency [Gurnett and Brice, 1966].

9.3.2.3 Instabilities

Since the first term in equation (9.3.42) always leads to damping, it is clear that any electromagnetic instabilities that occur for parallel propagation must arise from the second term, which is of the form

$$-\frac{k_{\parallel}}{|k_{\parallel}|} \int_0^{\infty} \left[v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} - v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] \pi v_{\perp}^2 dv_{\perp}. \quad (9.3.46)$$

To produce an instability, this term must be sufficiently positive to offset the damping caused by the first term. Since the integrand in the above equation is zero for an isotropic distribution, it is clear that an anisotropic distribution is required to drive an instability. The type of anisotropy required can be illustrated by evaluating (9.3.46) for a bi-Maxwellian velocity distribution function (??). The result for a bi-Maxwellian is easily shown to be

$$-\frac{k_{\parallel}}{|k_{\parallel}|} \int_0^{\infty} \left[v_{\parallel} \frac{\partial F_{s0}}{\partial v_{\perp}} - v_{\perp} \frac{\partial F_{s0}}{\partial v_{\parallel}} \right] \pi v_{\perp}^2 dv_{\perp} = \frac{k_{\parallel}}{|k_{\parallel}|} v_{\parallel} A_s \int_0^{\infty} F_{s0} 2\pi v_{\perp} dv_{\perp}, \quad (9.3.47)$$

where $A_s = (T_{s\perp}/T_{s\parallel} - 1)$ is a parameter called the anisotropy. A positive anisotropy corresponds to a disk-shaped velocity distribution, and a negative anisotropy corresponds to a cigar-shaped velocity distribution (see Figure 9.21). In principle, either type of anisotropy can lead to an instability. To determine the controlling factors, we substitute equation (9.3.47) into equation (9.3.42), which gives the following equation for the growth rate

$$\gamma = \pi \frac{1}{\partial D_r / \partial \omega} \sum_s \frac{\omega_{ps}^2}{\omega^2} \left[-\frac{\omega}{|k_{\parallel}|} - \frac{k_{\parallel}}{|k_{\parallel}|} v_{\parallel \text{Res}} A_s \right] \int_0^{\infty} F_{s0} 2\pi v_{\perp} dv_{\perp} \Bigg|_{v_{\parallel} = v_{\parallel \text{Res}}}. \quad (9.3.48)$$

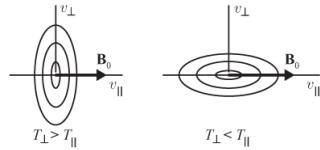


Fig. 9.21 A bi-Maxwellian distribution for two cases, $T_{\perp} > T_{\parallel}$ and $T_{\perp} < T_{\parallel}$.

This equation shows that the possible existence of an instability depends on the sign of

$k_{\parallel}v_{\parallel\text{Res}}A_s$. For an instability to occur, the product $k_{\parallel}v_{\parallel\text{Res}}A_s$ must be negative and sufficiently large to overcome the damping effect of the $-\omega/|k_{\parallel}|$ term. If $k_{\parallel}v_{\parallel\text{Res}}$ is positive, then instability can occur only if A_s is negative and, if $k_{\parallel}v_{\parallel\text{Res}}$ is negative, then instability can occur only if A_s is positive. To proceed further, it is necessary to examine some specific modes of propagation.

9.3.2.4 Whistler-mode instabilities

As an example of an electromagnetic instability driven by an anisotropy, we next consider the case of whistler-mode waves, for which ion motions can be ignored. Using the (+) sign in equation (9.3.37), and the high-density approximation introduced in Chapter ??, the real part of the dispersion relation for the whistler mode is given approximately by

$$D_r = -\frac{C^2 k_{\parallel}^2}{\omega^2} + \frac{\omega_p^2}{\omega(\omega_c - \omega)} = 0, \quad (9.3.49)$$

where we have neglected the 1 in equation (9.3.37) and used $\omega_p = \omega_{pe}$ and $\omega_c = \omega_{ce}$. Solving for the parallel wave number, the resonance velocity is given by

$$v_{\parallel\text{Res}} = -\frac{(\omega_c - \omega)}{k_{\parallel}} = -c \frac{k_{\parallel}}{|k_{\parallel}|} \frac{(\omega_c - \omega)^{3/2}}{\omega^{1/2} \omega_p}. \quad (9.3.50)$$

Since the frequency of the whistler mode is always less than the electron cyclotron frequency, the above equation shows that k_{\parallel} and v_{\parallel} must have opposite signs. The reason they must have opposite signs is that the wave frequency must be Doppler shifted up to the electron cyclotron frequency, i.e., $\omega' = \omega - k_{\parallel}v_{\parallel} = \omega_c$. Therefore, the guiding center of the resonant electrons moves in the opposite direction of the wave vector, as shown in Figure 9.22. Since $k_{\parallel}v_{\parallel\text{Res}}$ is negative, whistler-mode instabilities are always driven by a positive anisotropy in the electron velocity distribution, i.e., $T_{\perp}/T_{\parallel} > 1$ in the case of a bi-Maxwellian velocity distribution.

For resonance to occur, the wave field in the guiding center frame of the particle must be polarized in the same sense as the rotation of the resonant particle. For parallel propagation ($\theta = 0$), the wave field of the whistler mode rotates in the same right-hand sense as the electrons. The polarization condition is then clearly satisfied. Whistler-mode waves can also resonate with ions, which rotate in the opposite (left-hand) sense, assuming positively charged ions. However, the guiding center of the ion must then be moving in the same direction and faster than the wave, so that the polarization of the wave field is reversed (i.e., right-hand to left-hand) in the guiding center frame of the ion. Using the (+) sign in the general formula for the cyclotron resonance velocity, it is easy to show that the parallel resonance velocity for an ion is given by $v_{\parallel\text{Res}} = (\omega + \omega_{ci})/k_{\parallel}$, which confirms that the resonance velocity is greater than the phase velocity of the wave. This equation also shows that $k_{\parallel}v_{\parallel\text{Res}}$ is positive, so a negative ion anisotropy is required to drive the instability. Such ion-driven whistler-mode instabilities are rare, since the resonance velocities are usually very large, where there are few particles to resonate with the wave.

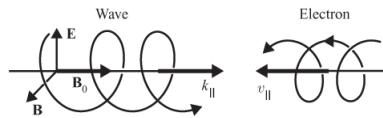


Fig. 9.22 For the whistler mode, the guiding center of resonant electrons moves in the opposite direction of the wave, i.e., $k_{\parallel}v_{\parallel} < 0$. Note that both the wave and the electrons are rotating in the same (right-hand) sense with respect to the zero-order magnetic field.

To estimate the number of particles in resonance with the wave, it is useful to compute the parallel resonant energy, which for electrons is given by

$$W_{\parallel\text{Res}} = \frac{1}{2}m_e v_{\parallel\text{Res}}^2. \quad (9.3.51)$$

Using equation (9.3.50), it is easy to show that the parallel resonant energy for the whistler mode is given to a good approximation by the equation

$$W_{\parallel\text{Res}} = W_c \left(1 - \frac{\omega}{\omega_c}\right)^3 \frac{\omega_c}{\omega}, \quad (9.3.52)$$

where

$$W_c = \frac{1}{2}mc^2 \left(\frac{\omega_c}{\omega_p}\right)^2 = \frac{B^2}{2\mu_0 n_0} \quad (9.3.53)$$

is a characteristic energy for whistler-mode resonance with electrons. Note that the characteristic resonant energy, $B^2/(2\mu_0 n_0)$, can be simply interpreted as the magnetic energy density per electron. For the characteristic energy to be non-relativistic, $W_c \ll (1/2)mc^2$, the electron plasma frequency must be substantially greater than the electron cyclotron frequency ($\omega_p \gg \omega_c$). Cyclotron resonance can also occur for relativistic energies, but the equations must be modified to account for relativistic effects. The frequency dependence of the parallel resonance energy is plotted in Figure 9.23 for the non-relativistic case. As can be seen, the resonance energy becomes very low near the cyclotron frequency, and very high at low frequencies. Since there are usually more electrons at low energies than at high energies, cyclotron resonant interactions are usually more important as the wave frequency approaches the cyclotron frequency.

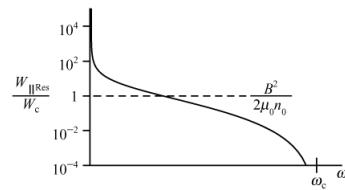


Fig. 9.23 The parallel resonance energy, W_{\parallel} , of the whistler mode as a function of frequency.

Having determined the resonance velocity, the growth rate of the whistler mode can be computed from

$$\gamma = \frac{-D_i}{\partial D_r / \partial \omega} = \pi \frac{k_{||}/|k_{||}|}{\partial D_r / \partial \omega} \frac{\omega_p^2}{\omega^2} G_0(v_{||\text{Res}}), \quad (9.3.54)$$

where $G_0(v_{||\text{Res}})$ is given by equation (9.3.41). Computing $\partial D_r / \partial \omega$ from equation (9.3.49) one obtains

$$\partial D_r / \partial \omega = \frac{\omega_p^2 \omega_c}{\omega^2 (\omega_c - \omega)^2}, \quad (9.3.55)$$

which, as previously commented, is always positive.

If the velocity distribution has a loss cone, as is often the case in a magnetic mirror machine or a planetary radiation belt, it is usually more convenient to express $F_0(\mathbf{v})$ as a function of the magnitude of the velocity, v , and the pitch angle, α , rather than as a function of the perpendicular velocity, v_{\perp} , and the parallel velocity, $v_{||}$. By transforming to spherical velocity space coordinates, it is easy to show that the integrand in the second term of equation (9.3.41) can be written in the following

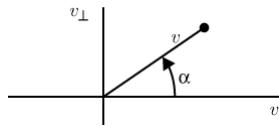


Fig. 9.24 The pitch angle, α , is the angle between the velocity vector and the $+v_{||}$ axis.

simple form

$$\left[v_{||} \frac{\partial F_0}{\partial v_{\perp}} - v_{\perp} \frac{\partial F_0}{\partial v_{||}} \right] = (\mathbf{v} \times \nabla_{\mathbf{v}} F_0) \phi = \frac{\partial F_0}{\partial \alpha}, \quad (9.3.56)$$

where the pitch angle α is measured from the $+v_{||}$ axis, as shown in Figure 9.24. Note that for an isotropic distribution, $\partial F_0 / \partial \alpha$ is zero, which is consistent with our earlier results concerning an isotropic distribution.

From equations (9.3.55) and (9.3.42) it can then be shown that the growth rate of the whistler mode is given by

$$\begin{aligned} \frac{\gamma}{\omega_c} &= \pi \left(1 - \frac{\omega}{\omega_c} \right)^2 \\ &\times \left[-\frac{\omega}{|k_{||}|} \int_0^\infty F_0 2\pi v_{\perp} dv_{\perp} + \frac{k_{||}}{|k_{||}|} \int_0^\infty (-\partial F_0 / \partial \alpha) \pi v_{\perp}^2 dv_{\perp} \right] \Bigg|_{v_{||}=v_{||\text{Res}}}, \end{aligned} \quad (9.3.57)$$

where $F_0(v, \alpha)$ is the electron distribution function. This equation can be put in a more convenient form by dividing by the first integral and substituting $v_{\perp} = |\tan \alpha| |v_{||\text{Res}}|$ for one of the v_{\perp} terms in the second integral. With these changes, equation (9.3.57) can be written in the form

$$\frac{\gamma}{\omega_c} = \pi \left(1 - \frac{\omega}{\omega_c}\right)^2 |v_{||\text{Res}}| \left(A - \frac{\omega}{|k_{||}v_{||\text{Res}}|}\right) \int_0^\infty F_0 2\pi v_\perp dv_\perp \Big|_{v_{||}=v_{||\text{Res}}} , \quad (9.3.58)$$

where the anisotropy A is now defined by

$$A = \frac{\int_0^\infty (k_{||}/|k_{||}|)(-\partial F_0/\partial \alpha) |\tan \alpha| \pi v_\perp dv_\perp}{\int_0^\infty F_0 2\pi v_\perp dv_\perp} \Big|_{v_{||}=v_{||\text{Res}}} , \quad (9.3.59)$$

and $|k_{||}|v_{||\text{Res}}|$ has been combined into one term $|k_{||}v_{||\text{Res}}|$.

The quantity $\omega/|k_{||}v_{||\text{Res}}|$ in equation (9.3.58) can be evaluated from the cyclotron resonance condition and is given by

$$\frac{\omega}{|k_{||}v_{||\text{Res}}|} = \frac{\omega}{\omega_c - \omega} . \quad (9.3.60)$$

The growth rate then becomes

$$\gamma = \pi \omega_c \left(1 - \frac{\omega}{\omega_c}\right)^2 |v_{||\text{Res}}| \left(A - \frac{\omega}{\omega_c - \omega}\right) \int_0^\infty F_0 2\pi v_\perp dv_\perp \Big|_{v_{||}=v_{||\text{Res}}} . \quad (9.3.61)$$

It is evident that the whistler mode is unstable whenever the anisotropy, A , is positive. The marginal stability condition ($\gamma = 0$) gives an upper frequency limit for the range of unstable frequencies and is given by

$$\omega^* = \frac{A}{1 + A} \omega_c . \quad (9.3.62)$$

At frequencies greater than ω^* , cyclotron damping, given by the first term in equation (9.3.42), is too strong for instability to occur. The growth rate in the unstable region depends strongly on the resonance velocity, which controls both the anisotropy and the number of particles in resonance with the wave.

It is easily demonstrated that a loss-cone type of distribution function has a positive anisotropy and is therefore unstable. As shown in Figure 9.25, for $k_{||} > 0$ the pitch angles involved in the resonant interaction occur in a region of velocity space where $\partial F/\partial \alpha < 0$, which from equation (9.3.59) implies a positive anisotropy, which is unstable. Electrons trapped in planetary radiation belts are often modeled by a distribution function of the form

$$F_0 \sim \frac{\sin^m \alpha}{v^n} . \quad (9.3.63)$$

For this type of distribution function, it is easy to show that the anisotropy factor is given by $A = m/2$. Note that a distribution function with a small m is nearly isotropic, whereas a distribution function with a large m is highly anisotropic.

Since electrons trapped in a planetary magnetic field always have a loss cone, whistler-mode instabilities of the type described above are a common feature of planetary magnetospheres, and play an important role in the pitch-angle scattering and loss of radiation belt

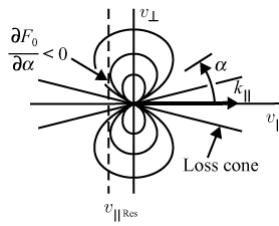


Fig. 9.25 A loss-cone electron velocity distribution showing the cyclotron resonance velocity, $v_{\parallel \text{Res}}$, for a whistler-mode wave propagating in the k_{\parallel} direction.

electrons [Kennel and Petschek, 1966]. Simple angular momentum considerations show that the generation of right-hand polarized electromagnetic angular momentum by growing whistler-mode waves must be associated with a reduction in the right-hand angular momentum of the electrons. The reduction in the angular momentum decreases the pitch angle and drives the resonant particles toward the loss cone, where they are lost from the system.

Two types of whistler-mode emissions occur in planetary radiation belts, hiss and chorus. Hiss has an almost featureless spectrum, and is believed to be generated by a marginally stable electron velocity distribution that arises from a balance between wave growth, which is driven by the loss-cone anisotropy, and pitch-angle scattering, which acts to scatter particles toward the loss cone, hence reducing the anisotropy. In contrast to hiss, chorus is highly structured and typically consists of many discrete tones, often rising in frequency, as shown in Figure 9.26. When played through a speaker, chorus is often described as sounding like the whistling songs that birds make when they wake up in the morning, called the “dawn chorus” in England. The complex frequency–time structure of chorus is believed to be caused by the trapping of resonant electrons in the rotating wave field. This trapping is similar in some respects to the electrostatic particle trapping discussed in Section ??, except the particles are trapped in the rotating electromagnetic wave field, rather than the longitudinal electrostatic field. For a numerical simulation of the non-linear trapping process believed to be responsible for the frequency–time structure of chorus, see Nunn *et al.* [1997], and for a discussion of the loss of radiation belt electrons due to interaction with these waves, see Horne and Thorne [2003].

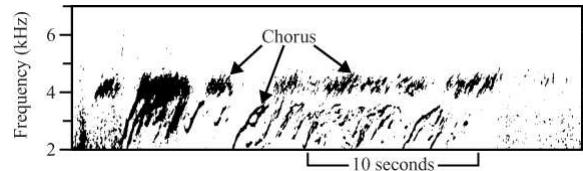


Fig. 9.26 A frequency–time spectrogram of whistler-mode chorus observed in the Earth’s magnetosphere.

9.3.2.5 Electromagnetic ion cyclotron mode instabilities

An instability very similar to the whistler-mode instability also occurs for electromagnetic ion cyclotron waves. In this case, the primary resonant interaction occurs between the left-hand polarized ion cyclotron wave and positively charged ions, which rotate in the left-hand sense relative to the magnetic field. For a two-component plasma consisting of electrons and one species of positive ions, it can be shown from equation (9.3.37) that the real part of the dispersion relation is given to a good approximation by

$$D_r = -\frac{c^2 k_{\parallel}^2}{\omega^2} + \frac{\omega_{pi}^2}{\omega_{ci}(\omega_{ci} - \omega)} = 0, \quad (9.3.64)$$

where we have neglected the 1 in equation (9.3.37) and made use of the fact that $\omega \ll |\omega_{ce}|$. This equation provides a good approximation for the index of refraction of the ion cyclotron mode over the frequency range $\omega \lesssim \omega_{ci}$, including the Alfvén wave regime near $\omega \approx 0$. Note that the functional form of the frequency dependence is somewhat different than for the whistler mode. This occurs because the electron term in equation (9.3.37) cannot be neglected for the ion cyclotron mode, whereas the ion term can be neglected for the whistler mode.

From equation (9.3.64), it is straightforward to show that the resonance velocity for the ion cyclotron mode is given by

$$v_{\parallel Res} = -\frac{(\omega_{ci} - \omega)}{k_{\parallel}} = -c \frac{k_{\parallel}}{|k_{\parallel}|} \frac{\omega_{ci}^{1/2} (\omega_{ci} - \omega)^{3/2}}{\omega \omega_{pi}}. \quad (9.3.65)$$

As with the whistler mode, $k_{\parallel} v_{\parallel Res}$ is negative, which implies that a positive (loss-cone type) anisotropy is required to drive the instability. Using the above equation it is easy to show that the parallel resonant energy is given by

$$W_{\parallel Res} = W_c \left(1 - \frac{\omega}{\omega_{ci}}\right)^3 \left(\frac{\omega_{ci}}{\omega}\right)^2, \quad (9.3.66)$$

where the characteristic energy is again defined by $W_c = B^2/(2\mu_0 n_0)$. Note that the only essential difference compared to the whistler mode is that the resonant energy increases more rapidly at low frequencies because of the ω^{-2} term.

After evaluating $\partial D_r / \partial \omega$ from the real part of the dispersion relation, which turns out to be

$$\frac{\partial D_r}{\partial \omega} = \frac{\omega_{pi}^2 (2\omega_{ci} - \omega)}{\omega \omega_{ci} (\omega_{ci} - \omega)^2}, \quad (9.3.67)$$

the growth rate can be written in the following form

$$\gamma = \pi \frac{(\omega_{ci} - \omega)^2 \omega_{ci}}{(2\omega_{ci} - \omega) \omega} |v_{\parallel Res}| \left(A - \frac{\omega}{(\omega_{ci} - \omega)}\right) \int_0^\infty F_{i0} 2\pi v_{\perp} dv_{\perp} \Big|_{v_{\parallel}=v_{\parallel Res}}, \quad (9.3.68)$$

where the ion anisotropy factor is defined in the same way as equation (9.3.59), except F_0 is now the ion distribution function, $F_{i0}(v, \alpha)$.

The electromagnetic ion cyclotron instability occurs whenever the ion anisotropy is sufficiently positive to offset the ion cyclotron damping term. Thus, the instability is driven by a loss-cone type of distribution, very similar to the situation with the whistler mode. Electromagnetic ion cyclotron waves have been observed in Earth's magnetosphere, and resonant interactions with these waves are believed to be one of the main mechanisms by which energetic ions are lost from planetary radiation belts. It is also worth noting that even small concentrations of minor ions (a few percent or less) can have important effects on the propagation and growth of ion cyclotron waves, see Summers and Thorne [2003].

9.3.3 *Oblique propagation*

It is evident from an inspection of the homogeneous equation (9.3.28) that the dispersion relation for oblique propagation is quite complicated. Each term in the matrix consists of an infinite sum of integrals involving Bessel functions. Also, there are no terms that are zero, such as occurred for parallel propagation. Since the dispersion relation, which is given by the determinant of the matrix in equation (9.3.28), has so many terms, it is not possible to obtain a general analytical solution for the roots of the dispersion relation, such as was possible for a cold magnetized plasma (see Chapter ??). Therefore, to obtain useful results, it is necessary either (1) to make very specific assumptions, or (2) to resort to numerical solutions. Because of the very restrictive nature of the assumptions needed to obtain usable analytical solutions, oblique propagation will not be discussed in detail. However, we can make some very general comments on the main features involved.

As one can see from the dielectric tensor elements in equations (9.3.19) through (9.3.27), all of the integrals involve a resonance denominator of the form $k_{\parallel}v_{\parallel} - \omega + n\omega_{cs}$. Therefore, the integrals can be evaluated using contour integration just as in the case of electrostatic waves. If one uses the weak growth rate approximation, as in Section ??, then the integrals split up into real and imaginary parts, similar to equations (9.2.34) and (9.2.35). The real part controls the propagation of the wave, i.e., $\omega(\mathbf{k})$, and the imaginary part controls the growth (or damping) of the wave, i.e., γ . For parallel propagation the Bessel function integrals are all zero, except the $n = 0$, and $n = \pm 1$ terms, and simply yield the cold plasma electromagnetic cyclotron modes. For oblique propagation, the higher order ($|n| > 1$) Bessel function integrals are, in general, non-zero, which introduces a series of new electromagnetic modes, one for each n , comparable in some respects to the electrostatic Bernstein modes. These modes have no counterpart in cold plasma theory. The numerical values of the integrals associated with these Bernstein-like electromagnetic modes are controlled by the $\beta_s = k_{\perp}v_{\perp}/\omega_{cs}$ term in the Bessel functions. If the thermal cyclotron radius is small compared to the perpendicular wavelength, such that $\beta_s \ll 1$, then these integrals are small and the corresponding modes are usually of negligible importance. On the other hand, if β_s is of order unity, these terms and their corresponding modes cannot be neglected. Sometimes, to isolate a specific mode, all the terms in the Bessel function series are omitted except for the one of interest. This approximation is usually valid only very close to the corresponding cyclotron harmonic, i.e., $\omega \sim n\omega_{cs}$.

Since the imaginary parts of the dielectric tensor elements control the growth rate, one can see that the $\partial F_0/\partial v_{\parallel}$ and $\partial F_0/\partial v_{\perp}$ terms provide the free energy source for any resonant instability that might occur. Therefore, the types of distribution functions that can give rise to wave growth for obliquely propagating electromagnetic waves are exactly the same as those that have already been discussed for electrostatic waves, namely beam and loss-cone types of velocity distributions. However, there is one difference. For electrostatic waves it is only the electric field that is involved in the resonant interaction, whereas for electromagnetic waves both the electric and the magnetic fields are involved. In fact, for some cases, such as the whistler mode in the high density limit, the wave magnetic field plays the dominant role. For a further discussion of oblique electromagnetic wave propagation in a hot magnetized plasma, see Kennel and Wong [1967] and Stix [1992].

Problems

9.1. The Harris dispersion relation is given by

$$\begin{aligned} D(k, \omega) = 1 - \sum_s \frac{\omega_{ps}^2}{k^2} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_0^{\infty} & \frac{J_n^2(\beta_s)}{k_{\parallel}v_{\parallel} - \omega + n\omega_{cs}} \\ & \times \left[k_{\parallel} \frac{\partial F_{0s}}{\partial v_{\parallel}} + \frac{n\omega_{cs}}{v_{\perp}} \frac{\partial F_{0s}}{\partial v_{\perp}} \right] 2\pi v_{\perp} dv_{\perp} dv_{\parallel} = 0, \end{aligned}$$

where $\beta_s = k_{\perp}v_{\perp}/\omega_{cs}$.

Show that the low-temperature, high-phase velocity limit of this equation is

$$D_0(k, \omega) = P \cos^2 \theta + S \sin^2 \theta = 0,$$

where $P = 1 - \sum_s (\omega_{ps}^2/\omega^2)$ and $S = 1 - \sum_s \omega_{ps}^2/(\omega^2 - \omega_{cs}^2)$ are the functions defined in cold plasma theory.

9.2. For the case of propagation perpendicular to the magnetic field ($k_{\parallel} = 0$) and a Maxwellian electron distribution, show that the Harris dispersion relation reduces to the following equation

$$D(k, \omega) = 1 - \sum_{n=1}^{\infty} \frac{2\omega_p^2}{\beta_c^2 \omega_c^2} \frac{e^{-\beta_c^2} I_n(\beta_c^2)}{(\omega/n\omega_c)^2 - 1} = 0$$

where $\beta_c = k_{\perp}\rho_c$, and $\rho_c = \sqrt{kT/m}/\omega_c$. The solutions to this dispersion relation are called the Bernstein modes.

9.3. Using a computer:

- (a) develop a computer program to solve for the roots of the Bernstein mode dispersion relation in problem 9.2 above;
- (b) plot the first four roots of $D(k_{\perp}, \omega) = 0$ as a function $\beta_c = k_{\perp}\rho_c$ for $\omega_p = (5/2)\omega_c$; this value for ω_p gives $\omega_{UHR} = 2.7\omega_c$;
- (c) compute the upper hybrid resonance frequency using the cold plasma equation

$\omega_{\text{UHR}} = \sqrt{\omega_c^2 + \omega_p^2}$; does your program give the proper value for ω_{UHR} in the limit $k_\perp \rho_c \rightarrow 0$;

- (d) find the ratio of the frequencies of the first three Q resonances to the electron cyclotron frequency (the Q resonances are the frequencies where $\partial\omega/\partial k_\perp = 0$).

9.4. Starting with the Harris dispersion relation and a Maxwellian distribution function

$$F_{s0} = \left(\frac{m_s}{2\pi\kappa T_s} \right)^{3/2} e^{-m_s(v_\perp^2 + v_\parallel^2)/2\kappa T_s},$$

show that the dispersion relation is given by

$$D(k, p) = 1 - \sum_{n,s} \frac{\Gamma_n(k_\perp \rho_{cs})^2}{(k \lambda_{Ds})^2} \left[\frac{1}{2} Z'(\zeta_s) - \frac{m \omega_{cs}}{k_\parallel} \sqrt{\frac{m_s}{2\kappa T_s}} Z(\zeta_s) \right] = 0$$

where

$$\begin{aligned} \Gamma_n(x) &= \exp[-x^2] I_n(x^2), \quad \rho_{cs} = \sqrt{\frac{\kappa T_s}{m_s}} \frac{1}{\omega_{cs}}, \quad \text{and} \\ \zeta_s &= \sqrt{\frac{m_s}{2\kappa T_s}} \frac{\omega - n\omega_{cs} + i\gamma}{k_\parallel}, \end{aligned}$$

and $\zeta_s = x_s + iy_s$ is given by

$$x_s = \sqrt{\frac{m_s}{2\kappa T_s}} \frac{\omega - n\omega_{cs}}{k_\parallel} \quad \text{and} \quad y_s = \sqrt{\frac{m_s}{2\kappa T_s}} \frac{\gamma}{k_\parallel}.$$

9.5. If we add a drift U_s to the parallel component of the Maxwellian, such that

$$F_{0s} = \left(\frac{m_s}{2\pi\kappa T_s} \right)^{3/2} e^{-m_s[(v_\parallel - U_s)^2 + v_\perp^2]/2\kappa T_s},$$

show that the resulting dispersion relation can be obtained from problem 9.4 above by making the substitution $\zeta_s \rightarrow (\zeta_s - \zeta_{U_s})$ where

$$\zeta_{U_s} = \sqrt{\frac{m_s}{2\kappa T_s}} U_s.$$

9.6. Show that the growth rate for electrostatic waves propagating in a two-component electron ion plasma with drifting Maxwellian distributions

$$F_{0s} = \left(\frac{m_s}{2\pi\kappa T_s} \right)^{3/2} e^{-m_s[(v_\parallel - U_s)^2 + v_\perp^2]/2\kappa T_s},$$

is given by

$$\frac{\gamma}{\omega_{ci}} = -\frac{k_\parallel}{|k_\parallel|} \sqrt{\frac{\pi}{2}} \sqrt{\frac{m_e}{\kappa T_e}} \left(\frac{T_e}{T_i} \right) \Gamma_1(k_\perp \rho_{ci}) \left(\frac{\omega}{k_\parallel} - U_e \right)$$

where

$$\Gamma_n(x) = \exp[-x^2] I_n(x^2), \quad \rho_{cs} = \sqrt{\frac{\kappa T_s}{m_s}} \frac{1}{\omega_{cs}}, \quad \text{and}$$

$$\zeta_s = \sqrt{\frac{m_s}{2\kappa T_s}} \frac{\omega - n\omega_{cs} + i\gamma}{k_{||}},$$

and $\zeta_{U_s} = \sqrt{m_s/2\kappa T_s} U_s$.

Determine the frequency, growth rate, and electron drift velocity threshold for the $n = 1$ current-driven ion cyclotron mode.

- 9.7. In the analysis of whistler-mode instabilities, it is convenient to approximate the distribution function in the resonance interaction region by a function of the form

$$F_0 = k \frac{\sin^m \alpha}{v^n}$$

where k is a constant, α is the pitch angle, and v is the magnitude of the velocity.

- (a) Make a sketch of contours of constant F as a function of α for $m = 0, m = 1$, and $m = 4$.
(b) Show that the anisotropy factor

$$A = \frac{\int_0^\infty (k_{||}/|k_{||}|)(-\partial F_0/\partial \alpha)| \tan \alpha | \pi v_\perp dv_\perp}{\int_0^\infty F_0 2\pi v_\perp dv_\perp} \Big|_{v_{||}=v_{||\text{Res}}}$$

is given by $A = m/2$ for this distribution function. Note, since $v_{||\text{Res}}$ is negative, α is in the range from $\pi/2$ to π , so $|\tan \alpha| = -\tan \alpha$.

- (c) Using the marginal stability condition $\omega^* = A\omega_c/(1+A)$ make a plot of the unstable frequency range as a function of m .

- 9.8. A distribution function consists of an anisotropic Maxwellian given by

$$F_0 = \left(\frac{m}{2\pi\kappa T_{||}} \right)^{1/2} \left(\frac{m}{2\pi\kappa T_\perp} \right) \exp \left[-\frac{mv_{||}^2}{2\kappa T_{||}} - \frac{mv_\perp^2}{2\kappa T_\perp} \right]$$

where $T_{||}$ and T_\perp are the parallel and perpendicular temperatures. Show that the anisotropy factor is

$$A = \frac{T_\perp}{T_{||}} - 1.$$

Again, to obtain the above result you must make use of the fact that $v_{||\text{Res}}$ is negative, so that $|\tan \alpha| = -\tan \alpha$.

- 9.9. Show that the index of refraction of the ion cyclotron mode is given to a good approximation by the equation

$$n^2 = \frac{\omega_{\text{pi}}^2}{\omega_{ci}(\omega_{ci} - \omega)},$$

and that the ion resonance energy is given by

$$W_{\text{Res}} = \frac{B^2}{2\mu_0 n_i} \left(1 - \frac{\omega}{\omega_{ci}}\right)^3 \left(\frac{\omega_{ci}}{\omega}\right)^2.$$

References

- Arfken, G. 1970. *Mathematical Methods for Physicists*. Academic Press: New York, pp. 488.
- Ashour-Abdalla, M., and Kennel, C. F. 1978. Nonconvective and convective electron cyclotron harmonic instabilities. *J. Geophys. Res.* **83**, 1531–1543.
- Bernstein, I. B. 1958. Waves in a plasma in a magnetic field. *Phys. Rev.* **109**, 10–21.
- Conkwright, N. B. 1934. *Differential Equations*. Macmillian: New York, pp. 19–25.
- Drummond, W. E., and Rosenbluth, M. N. 1962. Anomalous diffusion arising from microinstabilities in a plasma. *Phys. Fluids* **5**, 1507–1513.
- Gurnett, D. A., and Brice, N. M. 1966. Ion temperature in the ionosphere obtained from cyclotron damping of proton whistlers. *J. Geophys. Res.* **71**, 3639–3652.
- Harris, E. G. 1959. Unstable plasma oscillations in a magnetic field. *Phys. Rev. Lett.* **2**, 34–36.
- Horne, R. B. 1989. Path integrated growth of electrostatic waves: The generation of terrestrial myriametric radiation, *J. Geophys. Res.* **94**, 8895.
- Horne, R. B., and Thorne, R. M. 2003. Relativistic electron acceleration and precipitation during resonant interaction with whistler-mode chorus. *Geophys. Res. Lett.* **30** (10), 1527, doi: 10.1026/2003 GL 01697, 2003.
- Kennel, C. F., and Petschek, H. E. 1966. Limit on stably trapped particle fluxes. *J. Geophys. Res.* **71**, 1–28.
- Kennel, C. F., and Wong, H. V. 1967. Resonant particle instabilities in a uniform magnetic field. *J. Plasma Phys.* **1** (1), 75–80.
- Nunn, D., Omura, Y., Matsumoto, H., Nagano, I., and Yagitani, S. 1997. The numerical simulation of VLF chorus and discrete emissions observed on the Geotail spacecraft using a Vlasov code. *J. Geophys. Res.* **102** (A12), 27083–27097.
- Rönnmark, K. 1983. Computation of the dielectric tensor of a Maxwellian plasma. *Plasma Phys.* **25**, 699–701.
- Stix, T. H. 1992. *Waves in Plasmas*. American Institute of Physics: New York, pp. 237–293.
- Summers, D., and Thorne, R. M. 2003. Relativistic electron pitch-angle scattering by electromagnetic ion cyclotron waves during geomagnetic storms. *J. Geophys. Res.* **108** (A4), 1143, doi: 10.1029/2002 JA 009489, 2003.
- Tataronis, J. A., and Crawford, F. W. 1970. Cyclotron harmonic wave propagation and instabilities, I. Perpendicular propagation. *J. Plasma Phys.* **4**, 231–248.
- Watson, G. N. 1995. *A Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library Edition, Cambridge University Press: Cambridge, 1922, p. 395. Reprinted in 1995.

Further reading

- Chen, F. F. *Introduction to Plasma Physics and Controlled Fusion*. Plenum Press: New York, 1990, Chapter 7.
- Nicholson, D. R. *Introduction to Plasma Theory*. Krieger Publishing: Malabar, FL, 1983, Chapter 6.
- Stix, T. H. *Waves in Plasmas*. American Institute of Physics: New York, 1992, Chapters 10 and 11.
- Swanson, D. G. *Plasma Waves*. Academic Press: San Diego, CA, 1989, Section 4.3.

In this chapter we give a brief introduction to various types of non-linear effects that can occur in a plasma. Almost all of the basic equations in plasma physics have non-linear terms. For example, these include the $(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_v f$ term in the Vlasov equation (??), the $\mathbf{U} \cdot \nabla U$ term in the convective derivative (??), the $\rho_m \mathbf{U}$ term in the MHD mass continuity equation (??), and the $\mathbf{J} \times \mathbf{B}$ term in the MHD momentum equation (??). All of these terms represent potential sources of non-linear effects. There are many more. In our analysis of waves in the previous chapters, we always assumed that the wave amplitude was small, so that the governing equations can be linearized. This assumption provides a remarkably accurate description of many types of small amplitude waves. However, if the wave amplitude becomes large, as always occurs for an instability, the linearization assumption breaks down. Non-linear effects must then be taken into account. There are many such non-linear effects, more than we can possibly discuss in this introductory textbook. In order to limit the scope of the discussion, we will concentrate on two quite different types of non-linear analyses that have a wide range of applications. These are (1) quasi-linear theory, and (2) time-stationary electrostatic potentials. For a more comprehensive discussion of non-linear effects in plasmas, the reader is referred to one of the specialized books on non-linear effects, such as Kadomtsev [1965], Sagdeev and Galeev [1969], and Davidson [1972].

10.1 Quasi-linear theory

In quasi-linear theory, it is assumed that the plasma is only weakly unstable, and that the instability leads to a broad spectrum of waves that modifies the background plasma in a self-consistent way via non-linear interactions. To illustrate the basic features of quasi-linear theory, we limit our discussion to the temporal evolution of electrostatic waves in a one-dimensional unmagnetized collisionless plasma assumed to lie in the z direction. The governing equations for such a system are the Vlasov equation

$$\frac{\partial f_s}{\partial t} + v_z \frac{\partial f_s}{\partial z} - \frac{e_s}{m_s} \frac{\partial \Phi}{\partial z} \frac{\partial f_s}{\partial v_z} = 0, \quad (10.1.1)$$

and Poisson's equation

$$\frac{\partial^2 \Phi}{\partial z^2} = - \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} f_s \, dv_z, \quad (10.1.2)$$

where Φ is the electrostatic potential and f_s is the velocity distribution function of the s th species. We start by assuming that the plasma is initially in a homogeneous equilibrium state $f_s = f_{s0}(v_z)$ with no wave electric field, i.e., $\partial \Phi_0 / \partial z = 0$. It is easy to see that such an

equilibrium solution identically satisfies equations (10.1.1) and (10.1.2). The first term on the left of equation (10.1.1) vanishes because f_{s0} is time independent, the second term vanishes because f_{s0} is spatially uniform, and the last term vanishes because the equilibrium is field free. Equation (10.1.2) is satisfied because the plasma is quasi-neutral, i.e.,

$$\sum_s e_s \int_{-\infty}^{\infty} f_{s0} dv_z = 0. \quad (10.1.3)$$

At $t = 0$, we perturb the equilibrium by introducing a small perturbation $f_{s1}(z, v_z, t = 0)$ so that the total distribution function can be written as the sum of two terms

$$f_s(z, v_z, t = 0) = f_{s0}(v_z) + f_{s1}(z, v_z, t = 0). \quad (10.1.4)$$

Previously, in Chapter ??, we solved the problem by linearizing equations (10.1.1) and (10.1.2) about the equilibrium velocity distribution, $f_{s0}(v_z)$, and determining the time evolution of the first-order perturbation, $f_{s1}(z, v_z, t)$. In doing so, we assumed that the equilibrium distribution function $f_{s0}(v_z)$ at $t = 0$ remains unchanged as the perturbation evolves from its initial value $f_{s1}(z, v_z, t = 0)$ at $t = 0$ to its value $f_{s1}(z, v_z, t)$ at an arbitrary time t . Clearly, this assumption cannot be valid for all times. For instance, if the plasma is subject to a linear exponentially growing instability, f_{s1} will grow exponentially, invalidating the assumption that f_{s1} remain small. Then the question is how the growing $f_{s1}(z, v_z, t)$ will change the initial equilibrium $f_{s0}(v_z)$, and how that, in turn, will act back on $f_{s1}(z, v_z, t)$. As discussed below, the quasi-linear approximation provides one way to answer this question.

We start the discussion of quasi-linear theory by defining the spatial average of the time dependent distribution function $f_{s1}(z, v_z, t)$ as

$$\langle f_s \rangle(v_z, t) = \frac{1}{2L} \int_{-L}^L f_s(z, v_z, t) dz, \quad (10.1.5)$$

where the spatial integration is carried out over the entire length $2L$ of the one-dimensional plasma. We further require that the spatially averaged distribution function be identical to the zero-order distribution function at $t = 0$,

$$\langle f_s \rangle(v_z, t = 0) = f_{s0}(v_z), \quad (10.1.6)$$

but allow $\langle f_s \rangle$ to deviate from $f_{s0}(v_z)$ for $t > 0$. The total distribution function is then written as the sum of the averaged distribution function $\langle f_s \rangle$, and a fluctuation, i.e.,

$$f_s(z, v_z, t) = \langle f_s \rangle(v_z, t) + f_{s1}(z, v_z, t). \quad (10.1.7)$$

Note that at $t = 0$, equation (10.1.7) reduces identically to equation (10.1.4). The main difference between linear theory and quasi-linear theory lies in what one linearizes about: in linear theory, one linearizes about a time independent equilibrium distribution function $f_{s0}(v_z)$, whereas in quasi-linear theory one linearizes about a spatially averaged distribution function $\langle f_{s0} \rangle(v_z, t)$, that is allowed to vary slowly in time.

Taking the spatial average of both sides of equation (10.1.7) in the manner of equation

(10.1.5), it follows that the spatial average of the first-order perturbation of the velocity distribution function is zero

$$\langle f_{s1}(z, v_z, t) \rangle = 0. \quad (10.1.8)$$

For electrostatic perturbations, we can write the perturbed electric field as

$$E_1(z, t) = -\frac{\partial}{\partial z} \Phi_1(z, t). \quad (10.1.9)$$

Taking the spatial average of equation (10.1.9) and requiring that $\Phi_1(z, t)$ decays to zero at the system boundaries, we can show that the spatial average of the first-order electric field is also zero

$$\langle E_1(z, t) \rangle = -\frac{1}{2L} \int_{-L}^L \frac{\partial \Phi_1(z, t)}{\partial z} dz = 0. \quad (10.1.10)$$

10.1.1 The quasi-linear diffusion equation

Next, we develop an equation, called the quasi-linear diffusion equation, that describes the time evolution of the average distribution function. This equation is obtained by taking the spatial average of the Vlasov equation (10.1.1), which can be written

$$\frac{\partial}{\partial t} \langle f_s \rangle + \left\langle v_z \frac{\partial f_s}{\partial z} \right\rangle = \frac{e_s}{m_s} \left\langle \frac{\partial \Phi}{\partial z} \frac{\partial f_s}{\partial v_z} \right\rangle. \quad (10.1.11)$$

The second term on the left-hand side of the above equation can be rewritten as

$$\left\langle v_z \frac{\partial f_s}{\partial z} \right\rangle = \left\langle \frac{\partial}{\partial z} (v_z f_s) \right\rangle = 0, \quad (10.1.12)$$

because the integrand is a perfect differential and we require that the perturbation $f_{s1}(z, v_z, t)$ vanishes at the boundaries of the system. One way to ensure that this condition is satisfied is to take L to be very large and require that f_s decays to zero at large values of z . The term on the right-hand side of equation (10.1.11) can then be written

$$\frac{e_s}{m_s} \left\langle \frac{\partial \Phi}{\partial z} \frac{\partial f_s}{\partial v_z} \right\rangle = \frac{e_s}{m_s} \left\langle \left(\frac{\partial \Phi_0}{\partial z} + \frac{\partial \Phi_1}{\partial z} \right) \left(\frac{\partial \langle f_s \rangle}{\partial v_z} + \frac{\partial f_{s1}}{\partial v_z} \right) \right\rangle. \quad (10.1.13)$$

Since $\partial \Phi_0 / \partial z = 0$ in equilibrium and $\langle \partial \Phi_1 / \partial z \rangle = 0$ by equation (10.1.10), we obtain

$$\frac{e_s}{m_s} \left\langle \frac{\partial \Phi}{\partial z} \frac{\partial f_s}{\partial v_z} \right\rangle = \frac{e_s}{m_s} \left\langle \frac{\partial \Phi_1}{\partial z} \frac{\partial f_{s1}}{\partial v_z} \right\rangle = \frac{e_s}{m_s} \frac{\partial}{\partial v_z} \left\langle f_{s1} \frac{\partial \Phi_1}{\partial z} \right\rangle. \quad (10.1.14)$$

Using equations (10.1.12) and (10.1.14), we can reduce equation (10.1.11) to

$$\frac{\partial}{\partial t} \langle f_s \rangle = \frac{e_s}{m_s} \frac{\partial}{\partial v_z} \left\langle f_{s1} \frac{\partial \Phi_1}{\partial z} \right\rangle. \quad (10.1.15)$$

To derive an equation for $f_{s1}(z, v_z, t)$, we subtract the averaged equation (10.1.15) from the full Vlasov equation (10.1.1), which gives

$$\begin{aligned} \left(\frac{\partial}{\partial t} + v_z \frac{\partial}{\partial z} \right) f_{s1}(z, v_z, t) &= \frac{e_s}{m_s} \frac{\partial \Phi_1}{\partial z} \frac{\partial \langle f_s \rangle}{\partial v_z} \\ &+ \frac{e_s}{m_s} \frac{\partial}{\partial v_z} \left[\frac{\partial \Phi_1}{\partial z} f_{s1} - \left\langle \frac{\partial \Phi_1}{\partial z} f_{s1} \right\rangle \right]. \end{aligned} \quad (10.1.16)$$

Both equations (10.1.15) and (10.1.16) are exact. However, they are not a closed set because a time evolution equation for the second-order fluctuations $(\partial \Phi_1 / \partial z) f_{s1}$ (and its average) is needed. Any effort to determine such a time evolution equation will inevitably bring in third-order fluctuations. In turn, a time dependent equation for third-order fluctuations will involve fluctuations of the fourth order. This process, when continued, produces a hierarchy of indefinitely higher order. Therefore, there is a closure problem, similar to the closure problem encountered with the moment equations in Section ???. In order to truncate this hierarchy, it is necessary to introduce an approximation. The approximation used in quasi-linear theory is simply to neglect the last term on the right-hand side of equation (10.1.16). This approximation yields the equation

$$\left(\frac{\partial}{\partial t} + v_z \frac{\partial}{\partial z} \right) f_{s1}(z, v_z, t) = \frac{e_s}{m_s} \frac{\partial \Phi_1}{\partial z} \frac{\partial \langle f_s \rangle}{\partial v_z}. \quad (10.1.17)$$

We note that the above equation involves only fluctuating quantities of first order. The first-order electrostatic potential is obtained from Poisson's equation

$$\frac{\partial^2 \Phi_1}{\partial z^2} = - \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} f_{s1} dv_z, \quad (10.1.18)$$

which is exact. Equations (10.1.15), (10.1.17), and (10.1.18) now constitute a closed set of equations.

Equation (10.1.17) is formally similar to the time evolution equation for f_{s1} in linear theory except that the average distribution function $\langle f_s \rangle$ takes the place of the equilibrium distribution function f_{s0} . To solve these equations, we expand the spatial fluctuations, $\Phi_1(z, t)$ and $f_{s1}(z, v_z, t)$, in terms of Fourier transforms,

$$\Phi_1(z, t) = \int_{-\infty}^{\infty} \tilde{\Phi}_1(k, t) e^{ikz} dk, \quad (10.1.19)$$

and

$$f_{s1}(z, v_z, t) = \int_{-\infty}^{\infty} \tilde{f}_{s1}(k, v_z, t) e^{ikz} dk. \quad (10.1.20)$$

Substituting equations (10.1.19) and (10.1.20) into equations (10.1.17) and (10.1.18), and using the usual operator substitution $(\partial / \partial z \rightarrow ik)$ we obtain

$$\left(\frac{\partial}{\partial t} + ikv_z \right) \tilde{f}_{s1}(k, v_z, t) = \frac{e_s}{m_s} ik \tilde{\Phi}_1(k, t) \frac{\partial \langle f_s \rangle}{\partial v_z}, \quad (10.1.21)$$

and

$$k^2 \tilde{\Phi}_1(k, t) = \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} \tilde{f}_{s1}(k, v_z, t) dv_z. \quad (10.1.22)$$

To determine the time dependence of $\tilde{\Phi}_1(k, t)$ and $\tilde{f}_{s1}(k, v_z, t)$, we assume that these functions can be written in the form

$$\tilde{\Phi}_1(k, t) = \hat{\Phi}_1(k) e^{-i\omega(k,t)t}, \quad (10.1.23)$$

and

$$\tilde{f}_{s1}(k, v_z, t) = \hat{f}_{s1}(k, v_z) e^{-i\omega(k,t)t}, \quad (10.1.24)$$

where $\omega(k, t)$ is a complex frequency. If $\langle f_s \rangle$, in equation (10.1.21) were f_{s0} , then $\omega(k, t)$ would be independent of time. However, unlike f_{s0} , the average distribution function $\langle f_s \rangle$, changes as a function of time, in accordance with equation (10.1.15). Each realization of $\langle f_s \rangle$ in time produces its own frequency, ω , making ω a function of time as well. However, the time evolution of $\langle f_s \rangle$, and hence of ω , is slower than that of f_{s1} because, according to equation (10.1.15), $\partial \langle f_s \rangle / \partial t$ is second order in the fluctuations while $\partial f_{s1} / \partial t$ is first order. Hence, it is reasonable to assume that ω remains approximately constant on the faster time scale of fluctuations that evolve according to equations (10.1.21) and (10.1.22). Substituting equations (10.1.23) and (10.1.24) into equations (10.1.21) and (10.1.22), we obtain

$$\hat{f}_{s1}(k, v_z) = -\frac{e_s}{m_s} \frac{k \hat{\Phi}_1(k)}{\omega - kv_z} \frac{\partial \langle f_s \rangle}{\partial v_z} \quad (10.1.25)$$

and

$$k^2 \hat{\Phi}_1(k) = \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} \hat{f}_{s1}(k, v_z) dv_z. \quad (10.1.26)$$

We will assume that the wave is unstable, that is, the complex frequency ($\omega = \omega_r + i\gamma$) has a positive imaginary part ($\gamma > 0$). This makes a more detailed Landau analysis for damped modes, as discussed in Section ??, unnecessary. The generalization to include a Landau analysis, which covers stable as well as unstable modes, will not be carried out here.

Before we proceed to derive the quasi-linear diffusion equation, it is useful to introduce some identities involving Fourier representations. These follow from the requirement that the electrostatic potential must be real for all t , i.e.,

$$\Phi_1(z, t) = \Phi_1^*(z, t), \quad (10.1.27)$$

where $*$ denotes a complex conjugate. Written in terms of Fourier transforms, equation (10.1.27) becomes

$$\int_{-\infty}^{\infty} \hat{\Phi}_1(k) e^{ikz} e^{-i\omega^*(k,t)t} dk = \int_{-\infty}^{\infty} \hat{\Phi}_1^*(k) e^{-ikz} e^{i\omega^*(k,t)t} dk. \quad (10.1.28)$$

We note that k is a dummy variable of integration. So changing k to $-k$ will not change the value of the integral on the right-hand side. We can then rewrite the integral on the right-hand side as

$$\begin{aligned} \int_{-\infty}^{\infty} \hat{\Phi}_1^*(k) e^{-ikz} e^{i\omega(k,t)t} dk &= \int_{\infty}^{-\infty} \hat{\Phi}_1^*(-k) e^{ikz} e^{i\omega^*(-k,t)t} d(-k) \\ &= \int_{-\infty}^{\infty} \hat{\Phi}_1^*(-k) e^{ikz} e^{i\omega^*(-k,t)t} dk. \end{aligned} \quad (10.1.29)$$

Comparing the right-hand side of equation (10.1.29) with the left-hand side of equation (10.1.28), we obtain the identities

$$\hat{\Phi}_1(k) = \hat{\Phi}_1^*(-k), \quad (10.1.30)$$

and

$$\omega(k, t) = -\omega^*(-k, t). \quad (10.1.31)$$

Equation (10.1.31) implies that

$$\omega_r(k, t) = -\omega_r(-k, t), \quad \gamma(k, t) = \gamma(-k, t). \quad (10.1.32)$$

To obtain the quasi-linear diffusion equation from equation (10.1.15), consider the expression

$$\begin{aligned} \frac{\partial}{\partial v_z} \left\langle f_{s1} \frac{\partial \Phi_1}{\partial z} \right\rangle &= \frac{\partial}{\partial v_z} \frac{1}{2L} \int_{-L}^L \frac{\partial \Phi_1}{\partial z} f_{s1} dz \\ &= \frac{\partial}{\partial v_z} \frac{1}{2L} \int_{-L}^L \left[\left\{ \frac{\partial}{\partial z} \int_{-\infty}^{\infty} \tilde{\Phi}_1(k, t) e^{ikz} dk \right\} \right. \\ &\quad \times \left. \int_{-\infty}^{\infty} \tilde{f}_{s1}(k', v_z, t) e^{ik'z} dk' \right] dz, \end{aligned} \quad (10.1.33)$$

where we have used the Fourier transformation equations (10.1.19) and (10.1.20). If we now change the order of differentiation with respect to z and integration with respect to k in the first integrand (in parentheses), take the limit of a large box ($L \rightarrow \infty$), and use the identity

$$\lim_{L \rightarrow \infty} \int_{-L}^L e^{i(k+k')z} dz = 2\pi\delta(k+k'), \quad (10.1.34)$$

where $\delta(k+k')$ is the Dirac delta function, we obtain

$$\frac{\partial}{\partial v_z} \left\langle f_{s1} \frac{\partial \Phi_1}{\partial z} \right\rangle = -\frac{\pi}{L} \frac{\partial}{\partial v_z} \int_{-\infty}^{\infty} ik \tilde{\Phi}_1(-k, t) \tilde{f}_{s1}(k, v_z, t) dk. \quad (10.1.35)$$

Using equations (10.1.23) and (10.1.24) we can write the above equation as

$$\begin{aligned} \frac{\partial}{\partial v_z} \left\langle f_{s1} \frac{\partial \Phi_1}{\partial z} \right\rangle &= -\frac{\pi}{L} \frac{\partial}{\partial v_z} \int_{-\infty}^{\infty} ik \hat{\Phi}_1(-k) \hat{f}_{s1}(k, v_z) e^{-i[\omega(k,t)+\omega(-k,t)]t} dk \\ &= -\frac{\pi}{L} \frac{\partial}{\partial v_z} \int_{-\infty}^{\infty} ik \hat{\Phi}_1(-k) \hat{f}_{s1}(k, v_z) e^{2\gamma(k,t)t} dk, \end{aligned} \quad (10.1.36)$$

where we have used identities (10.1.32) to obtain

$$\begin{aligned} \omega(k, t) + \omega(-k, t) &= \omega_r(k, t) + i\gamma(k, t) + \omega_r(-k, t) + i\gamma(-k, t) \\ &= 2i\gamma(k, t). \end{aligned} \quad (10.1.37)$$

Substituting expression (10.1.25) for $\hat{f}_{s1}(k, v_z)$ into equation (10.1.36), we obtain

$$\frac{\partial}{\partial v_z} \left\langle f_{s1} \frac{\partial \Phi_1}{\partial z} \right\rangle = \frac{\pi}{L} \frac{e_s}{m_s} \frac{\partial}{\partial v_z} \int_{-\infty}^{\infty} ik^2 \frac{\hat{\Phi}_1(k) \hat{\Phi}_1(-k)}{\omega - kv_z} \frac{\partial}{\partial v_z} \langle f_s \rangle e^{2\gamma(k,t)t} dk. \quad (10.1.38)$$

Substituting the above equation into the right-hand side of equation (10.1.15), gives

$$\frac{\partial}{\partial t} \langle f_s \rangle = \frac{\pi}{L} \left(\frac{e_s}{m_s} \right)^2 \frac{\partial}{\partial v_z} \left[\int_{-\infty}^{\infty} \hat{E}_1(k) \hat{E}_1(-k) e^{2\gamma(k,t)t} \frac{i}{\omega - kv_z} \frac{\partial}{\partial v_z} \langle f_s \rangle \right] dk, \quad (10.1.39)$$

where $\hat{E}_1(k) = -ik\hat{\Phi}_1(k)$. The above equation can be written more compactly in terms of the averaged electrostatic energy, which is defined as

$$\langle W_E \rangle = \frac{\epsilon_0}{4L} \int_{-L}^L |E_1(z, t)|^2 dz, \quad (10.1.40)$$

where $\epsilon_0|E_1(z, t)|^2/2$ is the electrostatic energy density. Using the identity (10.1.34), it is easy to show that

$$\begin{aligned} \langle W_E \rangle &= \frac{\pi\epsilon_0}{2L} \int_{-\infty}^{\infty} \tilde{E}_1(k, t) \tilde{E}_1(-k, t) dk = \frac{\pi\epsilon_0}{2L} \int_{-\infty}^{\infty} \tilde{E}_1(k, t) \tilde{E}_1^*(k, t) dk \\ &= \frac{\pi\epsilon_0}{2L} \int_{-\infty}^{\infty} |\tilde{E}_1(k, t)|^2 dk \\ &= \int_{-\infty}^{\infty} \mathcal{E}(k, t) dk, \end{aligned} \quad (10.1.41)$$

where $\mathcal{E}(k, t) = (\pi\epsilon_0/2L)|\tilde{E}_1(k, t)|^2$ is called the *spectral density* of the electric field. We can also write $\mathcal{E}(k, t)$ in the equivalent form

$$\mathcal{E}(k, t) = \frac{\pi\epsilon_0}{2L} |\hat{E}_1(k)|^2 e^{2\gamma(k,t)t}. \quad (10.1.42)$$

The above equation implies that the spectral density $\mathcal{E}(k, t)$ obeys the differential equation

$$\frac{d\mathcal{E}(k, t)}{dt} = 2\gamma(k, t)\mathcal{E}(k, t), \quad (10.1.43)$$

where we recall that the time dependence of $\gamma(k, t)$ is suppressed in integrating $\mathcal{E}(k, t)$, because $\gamma(k, t)$ varies much more slowly with time than the fluctuating energy $\mathcal{E}(k, t)$. Using the identity $\hat{E}_1(-k) = \hat{E}_1^*(k)$ and equation (10.1.42), equation (10.1.39) can be written in the form of a diffusion equation

$$\frac{\partial}{\partial t}\langle f_s \rangle(v_z, t) = \frac{\partial}{\partial v_z} \left[D_q(v_z, t) \frac{\partial}{\partial v_z} \langle f_s \rangle(v_z, t) \right], \quad (10.1.44)$$

called the quasi-linear diffusion equation, where

$$D_q(v_z, t) = \frac{2}{\epsilon_0} \left(\frac{e_s}{m_s} \right)^2 \int_{-\infty}^{\infty} \frac{i\mathcal{E}(k, t)}{\omega - kv_z} dk, \quad (10.1.45)$$

is the quasi-linear diffusion coefficient. Since equation (10.1.44) is the time evolution equation for a real distribution function $\langle f_s \rangle(v_z, t)$, the coefficient $D_q(v_z, t)$, given by the above equation, must also be real. This can be seen by writing

$$\frac{i\mathcal{E}(k, t)}{\omega - kv_z} = \frac{i\mathcal{E}(k, t)}{\omega_r + i\gamma - kv_z} = \frac{i\mathcal{E}(k, t)[\omega_r - kv_z - i\gamma]}{(\omega_r - kv_z)^2 + \gamma^2}. \quad (10.1.46)$$

Using the fact that $\mathcal{E}(k, t)$ is even in k (i.e., $\mathcal{E}(k, t) = \mathcal{E}(-k, t)$) and $\omega_r(k, t)$ is odd in k (by identity (10.1.32)), it follows that the imaginary part of the integral in equation (10.1.45) vanishes identically because the integrand is odd in k . We are thus left with the coefficient

$$D_q(v_z, t) = \frac{2}{\epsilon_0} \left(\frac{e_s}{m_s} \right)^2 \int_{-\infty}^{\infty} \frac{\mathcal{E}(k, t)\gamma(k, t)}{[\omega_r(k, t) - kv_z]^2 + \gamma^2} dk, \quad (10.1.47)$$

which is manifestly real.

Equation (10.1.44), with the diffusion coefficient specified by equation (10.1.45) (or equation (10.1.47)) and \mathcal{E} given by equation (10.1.42), are the basic equations of quasi-linear theory.

10.1.2 Application to the bump-on-tail instability

To illustrate an application of quasi-linear theory, we next discuss the time evolution of the electron bump-on-tail instability. Such a distribution function is shown in Figure 10.1. The phase velocities, ω/k , of the unstable Langmuir waves that arise from this distribution function lie in the region where $f_{e0} = \langle f_e \rangle(v_z, t = 0)$ is an increasing function of v_z , i.e., where $\partial f_{e0}/\partial v_z > 0$. Assuming that the growth rate of the unstable waves is small, i.e., $|\gamma/\omega_r| \ll 1$, it can be shown that

$$\frac{\gamma}{(\omega_r - kv_z)^2 + \gamma^2} \simeq \pi\delta(\omega_r - kv_z). \quad (10.1.48)$$

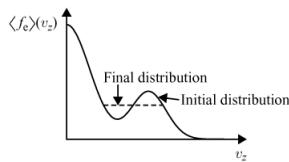


Fig. 10.1 The electrostatic instability caused by a double hump in the electron distribution function (solid line) leads to wave-particle interactions that act to flatten the distribution function (dashed line).

Substituting the above equation into the quasi-linear diffusion coefficient given by equation (10.1.47), we obtain

$$D_q(v_z, t) = \frac{2\pi}{\epsilon_0} \left(\frac{e}{m_e} \right)^2 \int_{-\infty}^{\infty} \mathcal{E}(k) \delta(\omega_r - kv_z) dk. \quad (10.1.49)$$

The quasi-linear diffusion equation (10.1.44) then becomes

$$\frac{\partial}{\partial t} \langle f_e \rangle(v_z, t) = \frac{\partial}{\partial v_z} \left\{ \frac{2\pi}{\epsilon_0 v_z} \left(\frac{e}{m_e} \right)^2 \mathcal{E} \left(k = \frac{\omega_r}{v_z} \right) \frac{\partial \langle f_e \rangle}{\partial v_z} (v_z, t) \right\}. \quad (10.1.50)$$

The quasi-linear diffusion coefficient in this case is non-zero in the region where the velocity v_z of the particles is resonant with the phase velocity ω_r/k of the waves, and is zero everywhere else in velocity space. Such a diffusion coefficient tends to smooth out velocity derivatives in the region of velocity space where $\partial f_{e0}/\partial v_z > 0$. However, this flattening process must also involve a larger region than the region where strictly $\partial f_{e0}/\partial v_z > 0$, because the particles removed from the unstable region, by particle conservation, must be taken up by the stable region nearby. The net outcome of the flattening process is the final distribution function shown in Figure 10.1. In the final state, the Langmuir waves are marginally stable, i.e., $\gamma(k, t \rightarrow \infty) = 0$ and $\mathcal{E}(k, t \rightarrow \infty)$ becomes stationary in time.

10.1.3 Chaotic velocity space diffusion

Quasi-linear theory predicts that an unstable particle distribution leads to the growth of a broad spectrum of waves, which in turn diffuses the particles in velocity space in such a way as to eliminate the region of positive slope that caused the instability. It is interesting to investigate exactly how this diffusion takes place. In Section ?? we discussed the trapping and phase mixing of particles in phase space when one wave is present. To investigate the non-linear effects that occur when a broad spectrum of waves is present, let us consider what happens when we superpose an infinite number of sinusoidal waves of the same amplitude and wavelength. For such a superposition, the equations of motion are

$$\frac{dz}{dt} = v_z, \quad (10.1.51)$$

and

$$m \frac{dv_z}{dt} = qE \sum_{n=-N}^N \cos(kz - n\omega t) \quad (10.1.52)$$

where $N \rightarrow \infty$. To simplify the equations we define dimensionless variables $\xi = kz$, $\tau = \omega t$, and $v = v_z/(\omega/k)$. The above equations can then be rewritten in the form

$$\frac{d\xi}{d\tau} = v, \quad (10.1.53)$$

and

$$\frac{dv}{d\tau} = \epsilon^2 \sum_{n=-N}^N \cos(\xi - n\tau), \quad (10.1.54)$$

where $\epsilon^2 = qkE/(m\omega^2)$ is a parameter that measures the depth of the potential well formed by a given wave. In the limit $N \rightarrow \infty$, it can be shown that the summation in the above equation can be written

$$\sum_{n=-N}^N \cos(\xi - n\tau) \xrightarrow{N \rightarrow \infty} 2\pi \cos \xi \sum_{n=-\infty}^{\infty} \delta(\tau - 2\pi n), \quad (10.1.55)$$

where $\delta(\tau - 2\pi n)$ is the Dirac delta function. In other words, the charged particle is subject to a sequence of instantaneous kicks when $\tau = 2\pi n$, which produces a series of discontinuous changes in its velocity. In between each kick, the particle moves with a uniform velocity. The process can be described rigorously by a series of difference equations, obtained by integrating the differential equations (10.1.53) and (10.1.54) with respect to time until just after the $(m+1)$ st kick. Defining a new variable $z_m = \xi_m/2\pi - 1/4$, it can be shown that equations (10.1.54) and (10.1.55) reduce to the so-called Standard Map, given by the difference equations

$$z_{m+1} = z_m + v_m, \quad (10.1.56)$$

and

$$v_{m+1} = v_m - 2\pi\epsilon^2 \sin(2\pi z_{m+1}). \quad (10.1.57)$$

The Standard Map, which is originally due to Chirikov [1969] and Taylor [1969], is one of the most widely studied examples of chaos in Hamiltonian systems, and its application to quasi-linear theory is discussed by Stix [1992]. Because this map is already in finite difference form, it can be integrated easily on a computer. The computing strategy is as follows. To start, choose a set of initial conditions (z_0, v_0) . For each initial condition, a series of iterants $(z_1, v_1), (z_2, v_2), \dots$, is then calculated and plotted as points in $z - v$ phase space. Figure 10.2 shows such a plot for $\epsilon = 0.10$. The v_n values cover a range suitable to show two of the infinite number of waves. We note the strong similarity to Figure ??, except that there are now trapping regions associated with each of the waves, only two of which are shown.

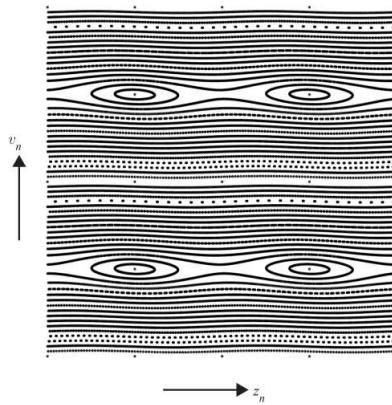


Fig. 10.2 Phase space of the Standard Map for $\epsilon = 0.10$.

As the wave amplitude increases, ϵ becomes larger, and the trapping regions increase in size. This trend can be seen in Figure 10.3, which is for $\epsilon = 0.30$. Well before the trapping regions become large enough to touch each other, we note the appearance of smaller trapping regions between the separatrices of the main trapping regions associated with the two waves. For an even larger wave amplitude, $\epsilon = 0.5$, shown in Figure 10.4, the phase space becomes a mixture of some regions that contain trapped particles and other regions that contain particles that appear to be randomly located. As ϵ becomes even larger, the trapping regions are destroyed and the particle trajectories are no longer constrained by nice, well-behaved functions that are constants of motion. Instead, the trajectories wander over and tend to fill most of the phase space, such as shown in Figure 10.5 for $\epsilon = 1.00$. Such behavior is characteristic of chaos, similar to the example discussed in Section ???. It is worth emphasizing again that the onset of chaos does *not* mean that the particle motions are random. Since the mapping is strictly deterministic, the final location of the particle in phase space is unique, no matter how large the value of ϵ may be. However, even small differences in the initial conditions or the wave amplitudes and phases can lead to very widely diverging final results when the particle motion is chaotic. In many situations of physical interest, such as a plasma comprising many particles, it is impossible to specify exactly the initial conditions of all the particles. The particle trajectories in phase space then become unpredictable, and hence random. This simple demonstration of chaotic particle motions provides a powerful justification for the basic assumption of quasi-linear theory, which is that the growth of waves over a range of phase velocities leads to the rapid diffusion of particles in velocity space and an eventual flattening of the distribution function in the unstable region of velocity space.

A good example of such velocity space diffusion can be seen in the spectra of solar flare electrons responsible for type III solar radio bursts. As discussed in Section 8.5.4.2, a type III radio burst is produced by mode conversion from Langmuir waves generated by energetic electrons from a solar flare. Because of time of flight considerations, a sharp velocity cutoff is introduced in the solar flare electron velocity distribution. This cutoff automatically leads to a region of positive slope, $\partial f_0 / \partial v_{\parallel} > 0$, and the growth of Langmuir

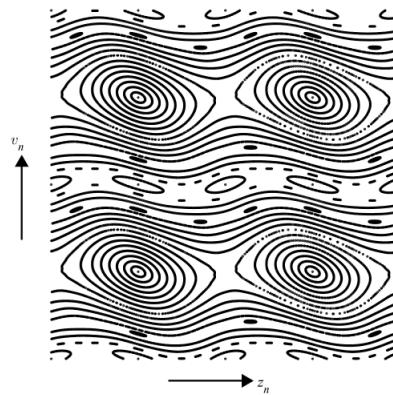


Fig. 10.3 Phase space of the Standard Map for $\epsilon = 0.30$.

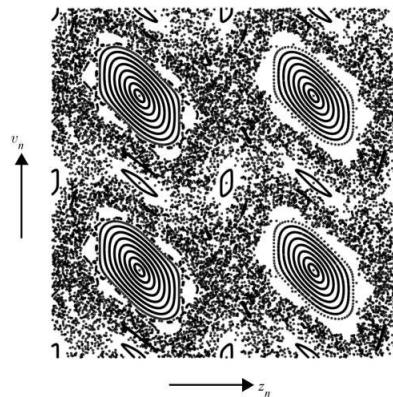


Fig. 10.4 Phase space of the Standard Map for $\epsilon = 0.50$.

waves, see Figure ???. An example of a series of electron velocity distribution functions observed in the solar wind near the orbit of the Earth following a solar flare is shown in Figure 10.6. This event produced strong Langmuir waves and an intense type III radio burst comparable to the example shown in Figure ???. Although the region of positive slope caused by the bump-on-tail distribution can be seen in the spectra at 0650 and 0705 UT, the velocity space cutoff is not nearly as sharp as would be expected from the time of flight cutoff (see Figure ??), and has been almost completely eliminated by the end of the event, at 0715 UT. This flattening of the electron velocity distribution function is caused by quasi-linear diffusion. In fact, regions with a substantial positive slope are seldom observed, since as soon as such a region occurs waves quickly grow that act to flatten the distribution function. Quasi-linear diffusion is also believed to play a crucial role in the propagation of solar flare electrons by limiting the amplitude of the Langmuir waves. In the absence of quasi-linear diffusion, Langmuir waves would grow to such large amplitudes that they would completely disrupt the bump-on-tail distribution, thereby preventing the electrons from propagating to great distances from the Sun, > 1 AU, where they are often observed.

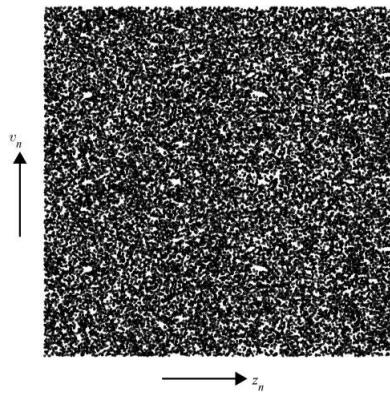


Fig. 10.5 Phase space of the Standard Map for $\epsilon = 1.00$.

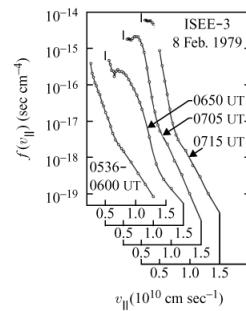


Fig. 10.6 A series of reduced one-dimensional electron velocity distribution functions showing the region of positive slope, $\partial f_0 / \partial v_{\parallel} > 0$, produced by electrons from a solar flare [from Lin et al., 1986]. The distribution labeled 0536-0600 UT is the pre-event solar wind distribution.

10.2 Stationary non-linear electrostatic potentials

As the second example of non-linear analysis techniques, we next consider a case in which the assumption of linearization breaks down completely and the full measure of the non-linearity inherent in the electrostatic Vlasov-Poisson system of equations (10.1.1) and (10.1.2) must be taken into account. As the general time dependent problem is very difficult and usually must be solved using numerical methods, we will focus our attention on time stationary ($\partial/\partial t = 0$) solutions of these equations. These solutions were first discovered by Bernstein, Greene, and Kruskal [1957] and are known as the BGK solutions.

In one dimension, the time stationary equations that must be solved are the Vlasov equation

$$v_z \frac{\partial f_s}{\partial z} - \frac{e_s}{m_s} \frac{\partial \Phi}{\partial z} \frac{\partial f_s}{\partial v_z} = 0, \quad (10.2.1)$$

and Poisson's equation

$$\frac{\partial^2 \Phi}{\partial z^2} = - \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} f_s \, dv_z. \quad (10.2.2)$$

As shown in Section ??, a general solution of the Vlasov equation for a time stationary electrostatic potential can be written in the form

$$f_s(v_z, z) = f_s(W_s), \quad (10.2.3)$$

where

$$W_s = \frac{1}{2} m_s v_z^2 + e_s \Phi(z), \quad (10.2.4)$$

is the total energy of a particle of type s . All that we now need to do is to solve Poisson's equation, which can be written

$$\frac{\partial^2 \Phi}{\partial z^2} = - \sum_s \frac{e_s}{\epsilon_0} \int_{-\infty}^{\infty} f_s \left(\frac{1}{2} m_s v_z^2 + e_s \Phi \right) \, dv_z. \quad (10.2.5)$$

If we assume that f_s is given, then the above equation becomes a differential equation for Φ . This is what we did in Section ?? when we specified that f_s is a Maxwellian, and with certain assumptions obtained the Debye-Hückel potential. Alternatively, if we assume that $\Phi(z)$ is given, then the above equation can be thought of as an integral equation for $f_s(W_s)$.

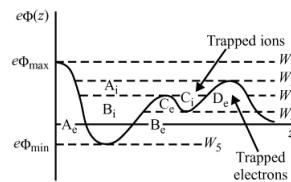


Fig. 10.7

An assumed time stationary electrostatic potential, $\Phi(z)$. In each region, bounded by total energies W_1, W_2, \dots , a unique solution can be found for any arbitrary $\Phi(z)$, provided only that $\Phi(z)$ be single valued and have a continuous second derivative.

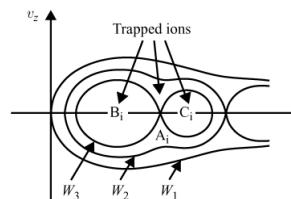


Fig. 10.8

The ion phase-space (z, v_z) boundaries for the first three total energy levels, W_1, W_2, W_3 , in Figure 10.7.

Next, we show the remarkable result that electron and ion distribution functions, f_e

and f_i , can be found that satisfy Poisson's equation and the Vlasov equation for any arbitrary $\Phi(z)$. The only requirements are that $\Phi(z)$ be single valued, have a continuous second derivative, and be finite. Consider the potential $e\Phi(z)$ shown in Figure 10.7. The particles moving in this potential can be classified according to their total energy, $W_s = (1/2)m_s v_z^2 + e_s \Phi(z)$. Untrapped electrons and ions occur at energies below $e\Phi_{\min}$ and above $e\Phi_{\max}$, respectively. The trapped particles can be subdivided by energy into separate trapping regions (A_e, B_e, C_e , and D_e for electrons, and A_i, B_i and C_i for ions) bounded by different turning points. Since the particles trapped in the various subregions cannot reach adjacent subregions, the distribution functions $f(W_s)$ can be independently prescribed in each subregion. This freedom to specify the distribution function independently in the various subregions can be illustrated more clearly by considering constant energy contours in $z - v_z$ phase space, as shown for the ions in Figure 10.8. Since the total energy is a constant of the motion, the phase-space trajectories of the particles are contours of constant energy. Since $d f_s/dt = 0$ along a representative particle trajectory, it follows that the distribution function is constant along a contour of constant total energy. Hence, once the distribution function is specified on a particular energy contour, it is determined at all points along the contour. However, in the trapping regions, the distribution functions can be independently specified because the energy contours are closed in the individual trapping regions and do not link to other regions of phase space.

This flexibility in specifying the distribution function for the electrons and ions in their respective trapping regions can be used to produce the charge density required to satisfy Poisson's equation for any specified $\Phi(z)$, the only requirement being that f_e and f_i are always positive. From equation (10.2.4), we obtain $dW_s = m_s v_z dv_z$, which implies that for all particles with $v_z \geq 0$, we have

$$f_s(v_z) dv_z = \frac{f_s(W_s) dW_s}{\sqrt{[2m_s(W_s - e_s \Phi)]}}. \quad (10.2.6)$$

To find the distribution function $f_s(W_s)$ giving any specified charge density, we must solve an integral equation of the form

$$\rho_{qs}(z) = e_s \int_{e_s \Phi}^{e_s \Phi_{\max}} \frac{f_s(W_s) dW_s}{\sqrt{2m_s(W_s - e_s \Phi)}}, \quad (10.2.7)$$

over a region where $\Phi(z)$ is monotonic, such as shown in Figure 10.9. In this equation Φ_{\max} is the maximum value of Φ in the region. The lower limit of integration is $e_s \Phi$ because no particle can have a total energy W_s less than $e_s \Phi$. Since $\Phi(z)$ is monotonic in any given subregion, $\rho_{qs}(z)$ can be specified as a function of $e_s \Phi$, i.e., $\rho_{qs}(z) = \rho_{s0}(e_s \Phi)$, so the above integral equation becomes

$$\rho_{s0}(e_s \Phi) = e_s \int_{e_s \Phi}^{e_s \Phi_{\max}} \frac{f_s(W_s) dW_s}{\sqrt{2m_s(W_s - e_s \Phi)}}. \quad (10.2.8)$$

We shall assume that $\rho_{qs}(a) = \rho_{s0}(e_s \Phi_{\max}) = 0$. It is easily verified that the following distribution function is an exact solution of the above integral equation

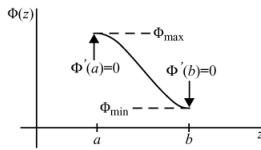


Fig. 10.9 A typical electrostatic potential profile between successive maxima and minima. For each such segment, a unique solution can be found for $f_s(v_z, z)$.

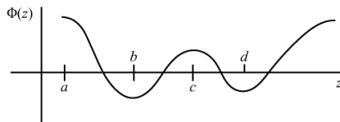


Fig. 10.10 To find a complete solution for $f_s(v_z, z)$, the electrostatic potential $\Phi(z)$ is broken up into segments, $a - b$, $b - c$, $c - d$, etc., between successive maxima and minima.

$$f_s(W_s) = -(2m_s)^{1/2} \frac{1}{e_s \pi} \int_{W_s}^{e_s \Phi_{\max}} dz \frac{d\rho_{s0}(z)}{dz} \frac{1}{\sqrt{z - W_s}}. \quad (10.2.9)$$

Using the above solution, the distribution function for the trapped particles can be determined in a piecewise fashion between successive maxima and minima of $\Phi(z)$ as illustrated in Figure 10.10. For example, between a and b , the ion distribution function, $f_i(W_i)$, can be determined over the range $e\Phi(a) > W_i > e\Phi(b)$, between b and c the electron distribution $f_e(W_e)$ can be determined over the range $-e\Phi(b) > W_e > -e\Phi(c)$, etc. Proceeding in this manner, a complete set of solutions for $f_i(v_z)$ and $f_e(v_z)$ can be obtained in all regions of the phase space for any given $\Phi(z)$.

10.2.1 Observations of BGK electrostatic potentials

The time stationary BGK electrostatic potentials described above represent a large class of solutions of the time stationary Vlasov and Poisson equations. Although these potentials are solutions of the time stationary equations, there is no guarantee that they are either stable or dynamically realizable in any given physical situation. Nevertheless, it appears that at least some of these solutions actually occur in nature.

For example, both laboratory and spacecraft measurements have shown that large amplitude time stationary electrostatic potentials, variously called “electrostatic shocks,” “double layers,” and “bipolar pulses,” can occur. Computer simulations have shown that these electrostatic structures are stable, and that they evolve from electrostatic plasma instabilities, such as the two-stream instability and the current-driven ion acoustic instability.

An example of a series of bipolar (i.e., positive as well as negative) electric field pulses observed in the Earth’s magnetosphere by the Geotail spacecraft is shown in Figure 10.11. Such bipolar pulses are commonly observed in regions where strong field aligned electron beams exist. Strong evidence now exists from numerical simulations that such bipolar pulses are the final state of an electron-beam driven instability. Figure 10.12 shows the

results of a computer simulation that self-consistently solves the non-linear evolution of an electron-beam driven instability by following the trajectories of individual particles in an electrostatic potential that is self-consistently determined from Poisson's equation. The computer simulation shows that in the final saturated state, the electric field has evolved into a series of nearly time independent bipolar pulses, shown in the top panel, which very closely resemble the electric field waveforms in Figure 10.11. The electrostatic potentials responsible for these fields appear to be closely related to the time stationary electrostatic solutions studied by Bernstein, Greene, and Kruskal [1957]. The cause of the bipolar pulses can be traced to the formation of "holes" in the electron phase-space distribution. A phase-space (v_z, z) diagram showing the formation of velocity space holes near the center of the bipolar pulses is shown in the bottom panel of Figure 10.12. In this case, the ions show almost no response, confirming that the bipolar pulses are primarily due to the formation of holes in the electron velocity distribution function.

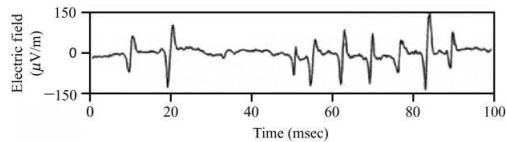


Fig. 10.11 A series of bipolar electric field pulses detected in the Earth's magnetotail by the Geotail spacecraft [from Matsumoto *et al.*, 1994].

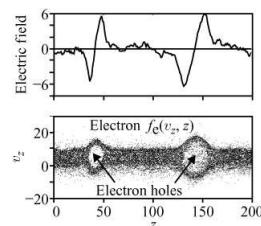


Fig. 10.12 The electric field, E , and electron distribution function $f_e(v_z, z)$ obtained from a computer simulation of an electron-beam driven two-stream instability [courtesy: L.-J. Chen].

We conclude by noting another application of the BGK solutions. It has been suggested recently [Hansen *et al.*, 1996] that the resolution of the anti-shielding paradox mentioned in Section ?? lies in recognizing that the assumption of a Maxwellian distribution, made in deriving the Debye length formula, is not, strictly speaking, correct. They note that a charged object in a plasma creates in its vicinity a potential well that traps particles of opposite charge. The distribution functions thus involve trapped as well as untrapped particles, and are better described as BGK solutions. The particles trapped in the vicinity of the charged object produce enhanced shielding, thus circumventing the anti-shielding mechanism discussed in Section ??.

Problems

10.1. Using the identity

$$\lim_{L \rightarrow \infty} \int_{-L}^L e^{i(k+k')z} dz = 2\pi\delta(k+k'),$$

Show that

$$\langle W_E \rangle = \frac{\pi\epsilon_0}{2L} \int_{-\infty}^{\infty} \tilde{E}_1(k, t) \tilde{E}_1(-k, t) dk = \int_{-\infty}^{\infty} \mathcal{E}(k, t) dk.$$

10.2. Show, by taking appropriate moments of the quasi-linear diffusion equation,

$$\frac{\partial}{\partial t} \langle f_s \rangle(v_z, t) = \frac{\partial}{\partial v_z} \left[D_q(v_z, t) \frac{\partial}{\partial v_z} \langle f_s \rangle(v_z, t) \right],$$

that particles, momentum, and energy are conserved.

10.3. Show in the limit of small γ/ω_r that the following equation reduces to a delta function

$$\frac{\gamma}{(\omega_r - kv_z)^2 + \gamma^2} \simeq \pi\delta(\omega_r - kv_z).$$

Hint: Consider the left-hand side as a function of ω_r and γ for fixed kv_z . For $\gamma \rightarrow 0$, integrate the left-hand side over ω_r to obtain the right-hand side.

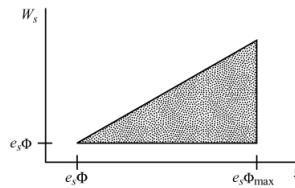
10.4. Show that the equation

$$f_s(W_s) = -(2m_s)^{1/2} \frac{1}{e_s \pi} \int_{W_s}^{e_s \Phi_{\max}} dz \frac{d\rho_{s0}(z)}{dz} \frac{1}{\sqrt{z - W_s}}$$

is a solution of the following integral equation.

$$\rho_{s0}(e_s \Phi) = e_s \int_{e_s \Phi}^{e_s \Phi_{\max}} \frac{f_s(W_s) dW_s}{\sqrt{2m_s(W_s - e_s \Phi)}}.$$

Hint: After substituting $\rho_{s0}(e_s \Phi)$ into the first equation, show that the resulting double integral involves an integration over the triangular area shown below.



Change the order of integration and make use of the following definite integral

$$\int_z^b \frac{dU}{\sqrt{(U-a)(b-U)}} = \pi.$$

References

- Bernstein, I. B., Greene, J. M., and Kruskal, M. D. 1957. Exact nonlinear plasma oscillations. *Phys. Rev.* **108**, 546.
- Chirikov, B. V. 1969. *Research Concerning the Theory of Nonlinear Resonances and Stochasticity*, Trans. A. T. Sanders, CERN Translation 71-40, Geneva, 1971. USSR Academy of Sciences: Novosibirsk, Report 267.
- Davidson, R. C. *Methods in Nonlinear Plasma Theory*. Academic Press: New York, 1972.
- Hansen, C., Reimann, A. B., and Fajans, J. 1996. Dynamic and Debye shielding and anti-shielding. *Phys. Plasmas* **3**, 1820.
- Kadomtsev, B. B. *Plasma Turbulence*. Academic Press: New York, 1965.
- Lin, R. P., Levedahl, W. K., Lotko, W., Gurnett, D. A., and Scarf, F. L. 1986. Evidence for nonlinear wave-wave interactions in solar type III radio bursts. *Astrophys. J.* **308**, 954–965.
- Matsumoto, H., Kojima, H., Miyatake, T., Omura, Y., Okada, M., Nagano, I., and Tsutsumi, M. 1994. Electrostatic solitary waves (ESW) in the magnetotail: BEN wave forms observed by Geotail. *Geophys. Res. Lett.* **21**, 2915–2918.
- Sagdeev, R. Z., and Galeev, A. A. *Nonlinear Plasma Theory*. Benjamin: New York, 1969.
- Taylor, J. B. 1969. Investigation of charged particle invariants. *Culham Lab. Prog. Report CLM-PR 12*, p. Th.12.

Further reading

- Chen, F. F. *Introduction to Plasma Physics and Controlled Fusion*. Plenum Press: New York, 1990, Chapter 8.
- Krall, N. A., and Trivelpiece, A. W. *Principles of Plasma Physics*. McGraw-Hill: New York, 1973, Chapter 10.
- Nicholson, D. R. *Introduction to Plasma Theory*. Krieger Publishing: Malabar, FL, 1983, Chapter 10.
- Schmidt, G. *Physics of High Temperature Plasmas*. Academic Press: New York, 1979, Chapter 9.
- Stix, T. H. *Waves in Plasmas*. American Institute of Physics: New York, 1992, Chapter 16.
- Swanson, D. G. *Plasma Waves*. Academic Press: San Diego, CA, 1989, Chapter 7.

Collisions between charged particles in a fully ionized plasma are mediated by the Coulomb force, which is a long-range force. By a long-range force, we mean a force that decays no faster than r^{-2} , where r is the distance between two charged particles. In contrast, collisions between charged and neutral particles are similar to those between hard spheres, which interact via a short-range force (Section ??). Due to the long-range Coulomb force, collective effects play a significant role in collisional processes. In this chapter, we will focus on collisional processes in a fully ionized plasma.

As discussed in Section ??, the number of electrons per Debye cube, $N_D = n_0 \lambda_D^3$, characterizes the discreteness of the plasma fluid. When N_D is much larger than one, the average kinetic energy is much larger than the average potential energy. Under these conditions, the average distance $\langle r \rangle$ between the nearest particles is much larger than the distance of closest approach, r_0 . This can be easily seen by rewriting N_D in the form,

$$N_D = \left(\frac{\langle r \rangle}{6\pi r_0} \right)^{3/2}, \quad (11.0.1)$$

where $\langle r \rangle = n_0^{-1/3}$ and $r_0 = e^2 / (6\pi\epsilon_0\kappa T)$. The distance of closest approach, r_0 , is calculated by equating the average kinetic energy, $(3/2)\kappa T$, to the potential energy, $e^2 / (4\pi\epsilon_0 r_0)$. When N_D is very large, particles rarely come as close as the distance of closest approach and therefore rarely suffer large-angle deflections. However, a large number of small-angle binary collisions can add up to produce a large deflection in the trajectory of the particle. If $N_D \gg 1$, it is possible to use $1/N_D$ as an expansion parameter in a statistical description that treats the collision process as the sum of a large number of small-angle collisions. This expansion fails when the condition $N_D \gg 1$ is no longer satisfied. Such plasmas are sometimes called strongly coupled plasmas [Ichimaru, 1994]. Strongly coupled plasmas occupy the region near and below the line $N_D = 1$ in Figure ??.

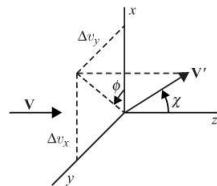


Fig. 11.1 The coordinate system used to analyze Coulomb collisions.

Before we develop a kinetic theory of collisional processes in a fully ionized plasma, it is useful to review the kinematics of binary Coulomb collisions which are the dominant form of collisions in such a plasma.

11.1 Binary Coulomb collisions

The scattering geometry of a charged particle of type s interacting with a charged particle of type s' is shown in Figure 11.1, as viewed from the center-of-mass frame. The vector \mathbf{V} is the relative velocity before the collision, and the vector \mathbf{V}' is the relative velocity after the collision. Because of conservation of energy the relative velocities before and after the collision are equal in magnitude, that is, $|\mathbf{V}| = |\mathbf{V}'|$. In the center-of-mass frame, the scattering angle χ and the impact parameter b obey the relation

$$\tan\left(\frac{\chi}{2}\right) = \frac{b_0}{b}, \quad (11.1.1)$$

where

$$b_0 = \frac{e_s e_{s'}}{4\pi\epsilon_0 \mu_{ss'} V^2}, \quad (11.1.2)$$

is the distance of closest approach, $V = |\mathbf{v} - \mathbf{v}'|$ is the relative velocity of the two particles at infinity, and

$$\mu_{ss'} = \frac{m_s m_{s'}}{m_s + m_{s'}} \quad (11.1.3)$$

is the reduced mass. In the case of electrons of mass m_e scattering off ions of mass m_i , where $m_i \gg m_e$, the reduced mass becomes $\mu_{ss'} \approx m_e$. The differential cross-section $\sigma_{ss'}(\chi)$, where χ is the scattering angle in the center-of-mass frame, is given by the Rutherford formula

$$\sigma_{ss'}(V, \chi) = \frac{1}{4} \frac{b_0^2}{\sin^4(\chi/2)}, \quad (11.1.4)$$

which generalizes equation (??) of Chapter ???. The total cross-section, σ_T , can be determined by integrating the differential cross-section over all solid angles, i.e.,

$$\begin{aligned} \sigma_T &= 2\pi \int_0^\pi \sigma_{ss'}(V, \chi) \sin \chi d\chi \\ &= \pi b_0^2 \int_0^\pi \frac{\sin(\chi/2) \cos(\chi/2)}{\sin^4(\chi/2)} d\chi \\ &= \pi b_0^2 \left[\frac{1}{\sin^2(\chi/2)} \right]_\pi^0, \end{aligned} \quad (11.1.5)$$

which diverges at the upper limit. It is clear by inspection of the above equation that this divergence arises due to scattering at very small angles (corresponding to very large impact parameters). Due to the long range of the Coulomb force, all particles are scattered, even those at very large distances, thereby causing the divergence.

The cross-section for scattering 90° or more in one binary collision is given by

$$\begin{aligned}\sigma_S &= 2\pi \int_{\pi/2}^{\pi} \sin \chi \sigma_{ss'}(V, \chi) d\chi = \frac{\pi b_0^2}{2} \int_{\pi/2}^{\pi} \sin \chi \frac{1}{\sin^4(\chi/2)} d\chi \\ &= \pi b_0^2.\end{aligned}\quad (11.1.6)$$

Unlike the total cross-section, this cross-section does not diverge because it does not include the contribution of particles with very large impact parameters (or very small scattering angles).

11.2 Importance of small-angle collisions

Next, let us consider an electron moving in a fully ionized plasma. The electron may suffer a large deflection, say 90° , either as a result of a single scattering event, or due to the cumulative effects of a large number of small-angle deflections that add up (in the sense of a random walk). We demonstrate below that the second process is much more important than the first.

Assume, without loss of generality, that the electron is moving initially with speed v in the z direction. As it moves through the plasma, the electron will suffer a sequence of random small-angle collisions. Let us assume that it suffers on the average N_c such collisions per unit length of its path projected along the z axis. For a plasma of average density n_0 , N_c is given by

$$N_c = n_0 \int_0^{b_m} 2\pi b db,\quad (11.2.1)$$

where b_m is the maximum impact parameter, to be estimated later. The parameter b_m is introduced in order to avoid the divergences associated with the Coulomb force, as discussed in Section 11.1. After N_c collisions, the electron acquires a transverse velocity ∇v_\perp with components

$$\Delta v_x = \sum_{i=1}^{N_c} \Delta v_{xi}, \quad \Delta v_y = \sum_{i=1}^{N_c} \Delta v_{yi},\quad (11.2.2)$$

where, as shown in Figure 11.1, we have

$$\Delta v_{xi} = V \sin \chi_i \cos \phi_i, \quad \Delta v_{yi} = V \sin \chi_i \sin \phi_i.\quad (11.2.3)$$

Here we have assumed, for simplicity, that the relative velocity in each collision is always in the z direction and that $\Delta v_{xi} = \Delta v_{yi} = 0$ before the collision. By inspection of equation (11.2.3), it is clear that because the scattering process is isotropic about the z axis, positive values of Δv_x and Δv_y are as likely as negative values. So if we take an average over many collisions per unit length of the particle path (denoted by the symbol $\langle \rangle$), we will obtain $\langle \Delta v_x \rangle = \langle \Delta v_y \rangle = 0$. However, the quantities $\langle (\Delta v_x)^2 \rangle$ and $\langle (\Delta v_y)^2 \rangle$ do not vanish in general.

For a single collision, we define $(\Delta v_{\perp})^2 = (\Delta v_x)^2 + (\Delta v_y)^2$. Using simple trigonometric identities and equation (11.1.1), we then obtain

$$\begin{aligned} (\Delta v_{\perp})^2 &= V^2 \sin^2 \chi = 4V^2 \sin^2 \frac{\chi}{2} \cos^2 \frac{\chi}{2} \\ &= \frac{4V^2(b/b_0)^2}{[1 + (b/b_0)^2]^2}. \end{aligned} \quad (11.2.4)$$

We can calculate $\langle(\Delta v_{\perp})^2\rangle$ per unit length, or $\langle(\Delta v_{\perp})^2\rangle/\Delta\ell$, by taking an average of $(\Delta v_{\perp})^2$ over all impact parameters and multiplying the result by N_c , given by equation (11.2.1). We obtain

$$\frac{\langle(\Delta v_{\perp})^2\rangle}{\Delta\ell} = N_c \frac{\int_0^{b_m} \Delta v_{\perp}^2 2\pi b \, db}{\int_0^{b_m} 2\pi b \, db} = n_0 \int_0^{b_m} \Delta v_{\perp}^2 2\pi b \, db. \quad (11.2.5)$$

Substituting expression (11.2.4) into equation (11.2.5) gives

$$\frac{\langle(\Delta v_{\perp})^2\rangle}{\Delta\ell} = 8\pi n_0 V^2 \int_0^{b_m} \frac{(b/b_0)^2}{[1 + (b/b_0)^2]^2} b \, db, \quad (11.2.6)$$

which reduces, for $b_m/b_0 \gg 1$, to

$$\frac{\langle(\Delta v_{\perp})^2\rangle}{\Delta\ell} \cong 8\pi n_0 V^2 b_0^2 \ln(b_m/b_0). \quad (11.2.7)$$

Since the long-range Coulomb force becomes exponentially small for impact parameters greater than the Debye length, the maximum impact parameter b_m is usually chosen to be the Debye length λ_D , see equation (??). The parameter b_0 , defined by equation (11.1.2), can be estimated by replacing $\mu_{ss'}V^2$ by its Maxwellian average, $3\kappa T$, see equation (??). With these substitutions it is then easy to show that the logarithmic term in equation (11.2.7) simplifies to

$$\begin{aligned} \ln \frac{b_m}{b_0} &= \ln \frac{\lambda_D}{b_0} \\ &= \ln \left(\frac{12\pi\epsilon_0\kappa T}{e^2} \lambda_D \right) = \ln(12\pi N_D) = \ln \Lambda, \end{aligned} \quad (11.2.8)$$

where the quantity in the parentheses occurs sufficiently often that it is defined to be a new parameter,

$$\Lambda = 12\pi N_D \quad (11.2.9)$$

called the plasma parameter. For most plasmas, $\ln \Lambda$ ranges from about 10 to 40.

Based on the above analysis, we can now estimate the mean-free path, λ_m , required for multiple small-angle collisions to produce a deflection of the order of 90° . Multiple small-angle collisions produce a change $\langle(\Delta v_{\perp})^2\rangle/\Delta\ell$, given by equation (11.2.7). To produce a deflection of the order of 90° in a path length, λ_m , we require that

$$\frac{\langle(\Delta v_{\perp})^2\rangle}{\Delta\ell} \lambda_m = V^2, \quad (11.2.10)$$

which, after substituting equation (11.2.7) for $\langle(\Delta v_{\perp})^2\rangle/\Delta\ell$, gives

$$\lambda_m = \frac{1}{8\pi n_0 b_0^2 \ln \Lambda}. \quad (11.2.11)$$

Next, we calculate the mean-free path for a single 90° large-angle collision, λ_s , which using equation (11.1.6), is given by

$$\lambda_s = \frac{1}{n_0 \sigma_s} = \frac{1}{n_0 \pi b_0^2}. \quad (11.2.12)$$

To compare the relative importance of a 90° deflection caused by a single collision to a 90° deflection produced by multiple small-angle collisions we take the ratio of the above equation to equation (11.2.11), which gives

$$\frac{\lambda_s}{\lambda_m} = 8 \ln \Lambda. \quad (11.2.13)$$

For a typical plasma $8 \ln \Lambda$ is of the order of 10^2 . Thus the likelihood that an electron will suffer a 90° deflection due to a single collision is much smaller than the likelihood that it will suffer the same deflection due to multiple small-angle collisions.

11.3 The Fokker–Planck equation

We have demonstrated that in a fully ionized plasma, small-angle collisions are much more probable than large-angle ones. Hence, the trajectory of a particle in a plasma is in general very complicated, composed of a very large number of seemingly random kicks which can eventually add up, in a root-mean-square sense, to a large deflection. To carry out a detailed analysis of such random motions, it is necessary to use a statistical analysis. For a statistical description it is useful to introduce a probability distribution function $P_s(\mathbf{v}, \Delta\mathbf{v})$ such that $P_s(\mathbf{v}, \Delta\mathbf{v}) d(\Delta\mathbf{v})$ is the probability that a test particle of type s with velocity \mathbf{v} at time t will change its velocity by $\Delta\mathbf{v}$ in time Δt as a result of collisions. In writing this distribution function, we have assumed that the probability distribution does not depend on time t explicitly, and furthermore, that it depends only on the particle velocity \mathbf{v} at time t and not on the particle velocity at previous times, that is, on the past history of the particle. In probability theory such a random process is called a stationary Markov process. We assume that the probability distribution function $P_s(\mathbf{v}, \Delta\mathbf{v})$ obeys the normalization condition

$$\int_{-\infty}^{\infty} P_s(\mathbf{v}, \Delta\mathbf{v}) d^3(\Delta\mathbf{v}) = 1. \quad (11.3.1)$$

Let $f_s(\mathbf{r}, \mathbf{v}, t)$ be the single-particle distribution function for particles of type s . If $f_s(\mathbf{r}, \mathbf{v} - \Delta\mathbf{v}, t - \Delta t)$ is the distribution function which evolves into the distribution function $f_s(\mathbf{r}, \mathbf{v}, t)$ due to collisions, by using the Markov distribution function $P_s(\mathbf{v}, \Delta\mathbf{v})$ defined above, we can write

$$f_s(\mathbf{r}, \mathbf{v}, t) = \int_{-\infty}^{\infty} f_s(\mathbf{r}, \mathbf{v} - \Delta\mathbf{v}, t - \Delta t) P_s(\mathbf{v} - \Delta\mathbf{v}, \Delta\mathbf{v}) d^3(\Delta\mathbf{v}). \quad (11.3.2)$$

We note that equation (11.3.2) assumes implicitly that the changes in the velocity dependence of the single-particle distribution function occur without a change in \mathbf{r} . In other words, we assume that a typical particle undergoes many collisions in time Δt without changing its spatial location significantly. Assuming that $|\Delta\mathbf{v}|$ is much smaller than $|\mathbf{v}|$ and Δt is much smaller than t , we can expand the integrand on the right-hand side of equation (11.3.2) to obtain

$$f_s(\mathbf{r}, \mathbf{v}, t) = \int_{-\infty}^{\infty} d^3(\Delta\mathbf{v}) \\ \times \left[\begin{array}{l} f_s(\mathbf{r}, \mathbf{v}, t) \mathbf{P}_s(\mathbf{v}, \Delta\mathbf{v}) - \Delta t \frac{\partial f_s(\mathbf{r}, \mathbf{v}, t)}{\partial t} \mathbf{P}_s(\mathbf{v}, \Delta\mathbf{v}) + \dots \\ - \Delta\mathbf{v} \cdot \nabla_{\mathbf{v}}(f_s(\mathbf{r}, \mathbf{v}, t)) \mathbf{p}_s(\mathbf{v}, \Delta\mathbf{v}) + \frac{1}{2} \Delta\mathbf{v} \Delta\mathbf{v} : \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} \{f_s(\mathbf{r}, \mathbf{v}, t)\} \\ \times \mathbf{P}_s(\mathbf{v}, \Delta\mathbf{v}) + \dots \end{array} \right] \quad (11.3.3)$$

where a single dot indicates the scalar product of two vectors, and double dots indicate the scalar product of two tensors which also produces a scalar. Using the normalization condition (11.3.1), we can now cancel the term on the left-hand side of the above equation with the first term on the right-hand side. If we divide the remaining terms in the above equation by Δt , take the limit $\Delta t \rightarrow 0$, and neglect the higher order terms in $\Delta\mathbf{v}$, we obtain

$$\frac{\delta_c f_s}{\delta t} = -\nabla_{\mathbf{v}} \cdot \left[\left\langle \frac{\Delta\mathbf{v}}{\Delta t} \right\rangle f_s \right] + \frac{1}{2} \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} : \left[\left\langle \frac{\Delta\mathbf{v} \Delta\mathbf{v}}{\Delta t} \right\rangle_s f_s \right], \quad (11.3.4)$$

where

$$\left\langle \frac{\Delta\mathbf{v}}{\Delta t} \right\rangle_s = \frac{1}{\Delta t} \int_{-\infty}^{\infty} \Delta\mathbf{v} P_s(\mathbf{v}, \Delta\mathbf{v}) d^3(\Delta\mathbf{v}), \quad (11.3.5)$$

is called the dynamical friction vector, and

$$\left\langle \frac{\Delta\mathbf{v} \Delta\mathbf{v}}{\Delta t} \right\rangle_s = \frac{1}{\Delta t} \int_{-\infty}^{\infty} \Delta\mathbf{v} \Delta\mathbf{v} P_s(\mathbf{v}, \Delta\mathbf{v}) d^3(\Delta\mathbf{v}), \quad (11.3.6)$$

is called the diffusion tensor. Equation (11.3.4) for $\delta_c f_s / \delta t$ is called the Fokker–Planck collision operator. Substituting $\delta_c f_s / \delta t$ into the Boltzmann equation (??) gives the Fokker–Planck equation

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{v}} f_s + \frac{\mathbf{F}}{m_s} \cdot \nabla_{\mathbf{v}} f_s = -\nabla_{\mathbf{v}} \cdot \left[\left\langle \frac{\Delta\mathbf{v}}{\Delta t} \right\rangle_s f_s \right] + \frac{1}{2} \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} : \left[\left\langle \frac{\Delta\mathbf{v} \Delta\mathbf{v}}{\Delta t} \right\rangle_s f_s \right]. \quad (11.3.7)$$

Although equation (11.3.7) is formally correct, it is incomplete because we have not yet determined the functional forms of the dynamical friction vector and the diffusion tensor. This is done in the next section.

11.3.1 The dynamical friction vector

Next we calculate the dynamical friction vector assuming that the collisional processes are dominated by binary Coulomb collisions, as discussed in Section 11.1. We start by considering collisions between a particle of type s moving with velocity \mathbf{v} and particles of type s' with velocities in d^3v' about \mathbf{v}' . If $\sigma_{ss'}$ represents the Rutherford scattering cross-section for collisions between a particle of type s and a particle of type s' , then $\sigma_{ss'}d\Omega f_{s'}(\mathbf{v}')d^3v'$ is the probability per unit length that a particle of type s suffers a collision with particles of type s' with velocities in d^3v' about \mathbf{v}' when the angle of scattering is in $d\Omega$ about Ω . By equation (11.1.4), we note that $\sigma_{ss'} = \sigma_{ss'}(V, \chi)$ where $V = |\mathbf{v} - \mathbf{v}'|$ is the relative velocity of two particles before the collision and χ is the angle of scattering in the center-of-mass frame. Hence, $V\sigma_{ss'}(V, \chi)d\Omega f_{s'}(\mathbf{v}')d^3v'$ is the probability per unit time that a collision will occur between a particle of type s moving with velocity \mathbf{v} and particles of type s' with velocities in d^3v' about \mathbf{v}' . Since the dynamical friction vector $\langle \Delta\mathbf{v}/\Delta t \rangle_s$ is the average change in the velocity of a particle of type s per unit time, we obtain

$$\left\langle \frac{\Delta\mathbf{v}}{\Delta t} \right\rangle_s = \sum_{s'} \int_{-\infty}^{\infty} d^3v' \int_0^{4\pi} V\sigma_{ss'}(V, \chi) f_{s'}(\mathbf{v}') \mathbf{v}' \Delta\mathbf{v} d\Omega, \quad (11.3.8)$$

where $\Delta\mathbf{v} = \mathbf{v} - \mathbf{v}_1$ and \mathbf{v}_1 is the velocity of the particle of type s after the collision. To proceed further we need to express $\Delta\mathbf{v}$ in terms of V and χ .

Let \mathbf{v}'_1 be the velocity of a particle of type s' after a collision. Defining the post-collision relative velocity $\mathbf{V}' = \mathbf{v}_1 - \mathbf{v}'_1$, we obtain the change in the relative velocity

$$\Delta\mathbf{V} = \mathbf{V}' - \mathbf{V} = (\mathbf{v}_1 - \mathbf{v}'_1) - (\mathbf{v} - \mathbf{v}') = \Delta\mathbf{v} - \Delta\mathbf{v}', \quad (11.3.9)$$

where $\Delta\mathbf{v} = \mathbf{v}_1 - \mathbf{v}$ and $\Delta\mathbf{v}' = \mathbf{v}'_1 - \mathbf{v}'$. By the law of conservation of linear momentum, it follows that

$$m_s \mathbf{v} + m_{s'} \mathbf{v}' = m_s \mathbf{v}_1 + m_{s'} \mathbf{v}'_1, \quad (11.3.10)$$

which gives

$$m_s \Delta\mathbf{v} + m_{s'} \Delta\mathbf{v}' = 0. \quad (11.3.11)$$

Eliminating $\Delta\mathbf{v}'$ between equations (11.3.9) and (11.3.11) gives

$$\Delta\mathbf{v} = \frac{\mu_{ss'}}{m_s} \Delta\mathbf{V}. \quad (11.3.12)$$

The vector $\Delta\mathbf{V}$ can be expressed in x , y and z coordinates as

$$\Delta \mathbf{V} = V \sin \chi \cos \phi \hat{\mathbf{x}} + V \sin \chi \sin \phi \hat{\mathbf{y}} - V(1 - \cos \chi) \hat{\mathbf{z}}, \quad (11.3.13)$$

where we have assumed $\mathbf{V} = V \hat{\mathbf{z}}$, as shown in Figure 11.1. Using equations (11.1.4), (11.3.12), and (11.3.13), we can evaluate the integral,

$$\int_0^{4\pi} \sigma_{ss'}(V, \chi) \Delta \mathbf{v} d\Omega = \frac{\mu_{ss'}}{m_s} \int_0^{4\pi} \sigma_{ss'}(V, \chi) \Delta \mathbf{V} d\Omega, \quad (11.3.14)$$

needed for the dynamical friction vector. We obtain

$$\begin{aligned} \int_0^{4\pi} \sigma_{ss'}(V, \chi) \Delta \mathbf{V} d\Omega &= \int_0^{2\pi} d\phi \int_{\chi_{\min}}^{\chi_{\max}} \sin \chi \sigma_{ss'}(V, \chi) \Delta \mathbf{V} d\chi \\ &= -2\pi V \hat{\mathbf{z}} \int_{\chi_{\min}}^{\chi_{\max}} \sin \chi (1 - \cos \chi) \sigma_{ss'}(V, \chi) d\chi \\ &= -2\pi \mathbf{V} \left(\frac{e_s e_{s'}}{4\pi \epsilon_0 \mu_{ss'} V^2} \right)^2 \int_{\chi_{\min}}^{\chi_{\max}} \frac{\sin \chi (1 - \cos \chi)}{4 \sin^4 \chi / 2} d\chi \\ &= -4\pi \mathbf{V} \left(\frac{e_s e_{s'}}{4\pi \epsilon_0 \mu_{ss'} V^2} \right)^2 \ln \Lambda_s, \end{aligned} \quad (11.3.15)$$

where χ_{\min} and χ_{\max} are the lower and upper cutoffs for the scattering angles and $\Lambda_s = \sin(\chi_{\max}/2)/\sin(\chi_{\min}/2)$. We will show below that $\ln \Lambda_s$ is closely related to $\ln \Lambda$, given by equation (11.2.9). Without introducing cutoffs, the integral (11.3.14) diverges logarithmically. Substituting equation (11.3.15) into equation (11.3.8), we obtain the dynamical friction vector

$$\left\langle \frac{\Delta \mathbf{v}}{\Delta t} \right\rangle_s = - \sum_{s'} 4\pi \ln \Lambda_s \left(\frac{e_s e_{s'}}{4\pi \epsilon_0 \mu_{ss'}} \right)^2 \frac{\mu_{ss'}}{m_s} \int_{-\infty}^{\infty} f_{s'}(\mathbf{v}') \frac{\mathbf{v} - \mathbf{v}'}{|\mathbf{v} - \mathbf{v}'|^3} d^3 v'. \quad (11.3.16)$$

To simplify the notation it is useful to define

$$\Gamma_{ss'} = \frac{e_s^2 e_{s'}^2}{4\pi \epsilon_0^2 m_s \mu_{ss'}} \ln \Lambda_s, \quad (11.3.17)$$

and

$$H_s(\mathbf{v}) = \int_{-\infty}^{\infty} \frac{f_s(\mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|} d^3 v', \quad (11.3.18)$$

and to use the identity

$$\nabla_{\mathbf{v}} H_s(\mathbf{v}) = - \int_{-\infty}^{\infty} \frac{\mathbf{v} - \mathbf{v}'}{|\mathbf{v} - \mathbf{v}'|^3} f_s(\mathbf{v}') d^3 v'. \quad (11.3.19)$$

With these definitions the dynamical friction vector can be written in the form

$$\left\langle \frac{\Delta \mathbf{v}}{\Delta t} \right\rangle_s = \sum_{s'} \Gamma_{ss'} \nabla_{\mathbf{v}} H_{s'}(\mathbf{v}). \quad (11.3.20)$$

The function $H_{s'}(\mathbf{v})$, defined by equation (11.3.18), is called the first Rosenbluth potential [Rosenbluth *et al.*, 1957]. Its mathematical form is analogous to the electrostatic potential $\Phi(\mathbf{r})$ due to a distribution of charge $\rho_q(\mathbf{r})$,

$$\Phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int_{-\infty}^{\infty} \frac{\rho_q(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\chi', \quad (11.3.21)$$

which produces the electric field, $\mathbf{E} = -\nabla\Phi$. This analogy leads to some useful qualitative conclusions. For instance, if a charge distribution is spherically symmetric, then $\mathbf{E} = E(r)\hat{\mathbf{r}}$. Furthermore, \mathbf{E} at any point \mathbf{r} depends only on the charge contained *within* the sphere centered at the origin and of radius r , not on the charge outside the sphere. Analogously, if $f_s(\mathbf{v})$ is a spherically symmetric distribution function in velocity space, that is, $f_s(\mathbf{v}) = f_s(v)$, then the dynamical friction vector on a fast particle of velocity \mathbf{v} must depend only on \mathbf{v} and on particles moving *slower* than the fast particle, not on particles moving faster than the fast particle. By equation (11.3.19), the direction of this force is such as to slow down the fast particle, which explains why it is called a frictional force.

11.3.2 The diffusion tensor

Using the same approach used in the last section, the diffusion tensor (11.3.6) can be written

$$\left\langle \frac{\Delta \mathbf{v} \Delta \mathbf{v}}{\Delta t} \right\rangle_s = \sum_{s'} \left(\frac{\mu_{ss'}}{m_s} \right)^2 \int_{-\infty}^{\infty} d^3 v' \int_0^{4\pi} V \sigma_{ss'}(V, \chi) f_{s'}(\mathbf{v}') \Delta \mathbf{v} \Delta \mathbf{v} d\Omega. \quad (11.3.22)$$

In the above equation it is useful to define the following tensor integral, which using equation (11.3.13) can be written

$$\begin{aligned} \overset{\leftrightarrow}{\mathbf{I}} &= \int_0^{4\pi} \sigma_{ss'}(V, \chi) \Delta \mathbf{v} \Delta \mathbf{v} d\Omega \\ &= 2\pi V^2 \int_{\chi_{\min}}^{\chi_{\max}} \sin \chi \sigma_{ss'}(V, \chi) \left[(\hat{\mathbf{x}}\hat{\mathbf{x}} + \hat{\mathbf{y}}\hat{\mathbf{y}}) \frac{\sin^2 \chi}{2} + \hat{\mathbf{z}}\hat{\mathbf{z}} (1 - \cos \chi)^2 \right] d\chi. \end{aligned} \quad (11.3.23)$$

It follows from equation (11.3.23) that $\overset{\leftrightarrow}{\mathbf{I}}$ is diagonal with

$$I_{xx} = I_{yy} = \pi V^2 \int_{\chi_{\min}}^{\chi_{\max}} \sin^3 \chi \sigma_{ss'}(V, \chi) d\chi, \quad (11.3.24)$$

and

$$I_{zz} = 2\pi V^2 \int_{\chi_{\min}}^{\chi_{\max}} \sin \chi (1 - \cos \chi)^2 \sigma_{ss'}(V, \chi) d\chi. \quad (11.3.25)$$

Substituting expression (11.1.4) for $\sigma_{ss'}$ into equations (11.3.24) and (11.3.25) and carrying out the integrals over χ , we obtain

$$I_{xx} = I_{yy} = \frac{1}{4\pi} \left(\frac{e_s e_{s'}}{\epsilon_0 \mu_{ss'} V} \right)^2 \left(\ln \Lambda_s + \frac{\cos \chi_{\max}}{4} - \frac{\cos \chi_{\min}}{4} \right). \quad (11.3.26)$$

As discussed below, $\ln \Lambda_s$ is typically much larger than unity, so we can approximate I_{xx} and I_{yy} by

$$I_{xx} = I_{yy} \simeq \frac{1}{4\pi} \left(\frac{e_s e_{s'}}{\epsilon_0 \mu_{ss'} V} \right)^2 \ln \Lambda_s. \quad (11.3.27)$$

Similarly, equation (11.3.25) yields

$$I_{zz} = \frac{1}{8\pi} \left(\frac{e_s e_{s'}}{\epsilon_0 \mu_{ss'} V} \right)^2 (\cos \chi_{\min} - \cos \chi_{\max}). \quad (11.3.28)$$

It is clear by inspection of equations (11.3.27) and (11.3.28) that $I_{xx} = I_{yy} \gg I_{zz}$ because $\ln \Lambda_s \gg 1$. As we have neglected terms of order I_{zz} in reducing equation (11.3.26) to (11.3.27), it is consistent to assume that

$$I_{zz} \simeq 0. \quad (11.3.29)$$

Equations (11.3.27) and (11.3.29) have been derived assuming a coordinate system in which $\mathbf{V} = V \hat{\mathbf{z}}$. It is easy to see that the coordinate independent generalization of $\overset{\leftrightarrow}{\mathbf{I}}$ is given by

$$\overset{\leftrightarrow}{\mathbf{I}} \simeq \frac{1}{4\pi} \left(\frac{e_s e_{s'}}{\epsilon_0 \mu_{ss'} V} \right)^2 \left[\overset{\leftrightarrow}{\mathbf{1}} - \frac{\mathbf{V}\mathbf{V}}{V^2} \right] \ln \Lambda_s, \quad (11.3.30)$$

where $\overset{\leftrightarrow}{\mathbf{1}}$ is the unit tensor. Substituting the integral (11.3.30) into equation (11.3.22), we obtain the diffusion tensor

$$\left\langle \frac{\Delta \mathbf{v} \Delta \mathbf{v}}{\Delta t} \right\rangle_s = \sum_{s'} \Gamma_{ss'} \frac{\mu_{ss'}}{m_s} \int_{-\infty}^{\infty} \frac{f_{s'}(\mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|} \left[\overset{\leftrightarrow}{\mathbf{1}} - \frac{(\mathbf{v} - \mathbf{v}')(\mathbf{v} - \mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|^2} \right] d^3 v', \quad (11.3.31)$$

where we recall that $\mathbf{V} = \mathbf{v} - \mathbf{v}'$. Equation (11.3.31) can be simplified further by using the identity

$$\nabla_{\mathbf{v}} \nabla_{\mathbf{v}} V = \nabla_{\mathbf{v}} \left(\frac{\mathbf{V}}{V} \right) = \frac{1}{V} \left(\overset{\leftrightarrow}{\mathbf{1}} - \frac{\mathbf{V}\mathbf{V}}{V^2} \right), \quad (11.3.32)$$

and the definition

$$G_s(\mathbf{v}) = \int_{-\infty}^{\infty} f_s(\mathbf{v}') |\mathbf{v} - \mathbf{v}'| d^3 v', \quad (11.3.33)$$

whereupon we obtain

$$\left\langle \frac{\Delta \mathbf{v} \Delta \mathbf{v}}{\Delta t} \right\rangle_s = \sum_{s'} \Gamma_{ss'} \frac{\mu_{ss'}}{m_s} \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} G_{s'}(\mathbf{v}). \quad (11.3.34)$$

The function $G_s(\mathbf{v})$ is known as the second Rosenbluth potential [Rosenbluth *et al.*, 1957]. It is related to the first Rosenbluth potential $H_s(\mathbf{v})$ by the identity

$$H_s(\mathbf{v}) = \frac{1}{2} \nabla_{\mathbf{v}} \cdot \nabla_{\mathbf{v}} G_s(\mathbf{v}). \quad (11.3.35)$$

Substituting the dynamical friction vector (11.3.20) and the diffusion tensor (11.3.34) into equation (11.3.4) for the Fokker–Planck collision operator, gives

$$\frac{\delta_c f_s}{\delta t} = \sum_{s'} \Gamma_{ss'} \left[-\nabla_{\mathbf{v}} \cdot (f_s \nabla_{\mathbf{v}} H_{s'}) + \frac{1}{2} \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} : (f_s \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} G_{s'}) \right]. \quad (11.3.36)$$

It remains to discuss the parameter

$$\ln \Lambda_s = \ln \frac{\sin(\chi_{\max}/2)}{\sin(\chi_{\min}/2)}. \quad (11.3.37)$$

Let us denote the impact parameters that produce the scattering angles χ_{\min} and χ_{\max} by b_m and b_{\min} , respectively. Using relation (11.1.1), we obtain

$$\sin^2 \frac{\chi}{2} = \frac{b_0^2}{b_0^2 + b^2}, \quad (11.3.38)$$

which implies that

$$\ln \Lambda_s = \frac{1}{2} \ln \left(\frac{b_0^2 + b_m^2}{b_0^2 + b_{\min}^2} \right). \quad (11.3.39)$$

As discussed in Section 11.2, the parameter b_m is usually chosen to be the Debye length λ_D , while the parameter b_0 is usually taken to be given by $b_0 = e_s e_s' / (12\pi\epsilon_0\kappa T)$ where T is the temperature of the plasma (assuming the species have attained thermal equilibrium). Since there is some freedom in the choice of b_{\min} , one may choose either $b_{\min} = b_0$ or $b_{\min} = 0$. Either choice leads to values of $\ln \Lambda_s$ that differ from each other to the extent of $\ln \sqrt{2}$ which is of order unity. To be consistent with Section 11.2 we choose $b_{\min} = 0$. Since $b_m \gg b_0$, we then obtain, by equation (11.2.8),

$$\ln \Lambda_s \simeq \ln \left(\frac{b_m}{b_0} \right) = \ln(12\pi N_D) = \ln \Lambda. \quad (11.3.40)$$

11.4 Conductivity of a fully ionized plasma

As an application of the Fokker–Planck equation, we next compute the conductivity of a plasma using the Lorentz gas approximation. In this approximation, electrons are assumed to collide with fixed immobile ions. Since the ions are immobile (i.e., $m_i \gg m_e$), the ion distribution function can be represented by a Dirac delta function, that is

$$f_i(\mathbf{v}_i) = n_0 \delta(\mathbf{v}_i), \quad (11.4.1)$$

where n_0 is the number density of electrons and ions. Substituting equation (11.4.1) into equations (11.3.18) and (11.3.33), we obtain the Rosenbluth potentials

$$H_i(\mathbf{v}) \simeq \frac{n_0}{v}, \quad (11.4.2)$$

and

$$G_i(\mathbf{v}) \simeq n_0 v. \quad (11.4.3)$$

The Fokker–Planck collision operator then becomes

$$\frac{\delta_c f_e}{\delta t} \simeq n_0 \Gamma_{ei} \left[-\nabla_{\mathbf{v}} \cdot \left(f_e \nabla_{\mathbf{v}} \left(\frac{1}{v} \right) \right) + \frac{1}{2} \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} : (f_e \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} v) \right], \quad (11.4.4)$$

where, by equation (11.3.17),

$$\Gamma_{ei} = \frac{e^4}{4\pi\epsilon_0^2 m_e^2} \ln \Lambda. \quad (11.4.5)$$

We use the identities

$$\nabla_{\mathbf{v}} \left(\frac{1}{v} \right) = -\frac{1}{v^2} \nabla_{\mathbf{v}} V = -\frac{\mathbf{v}}{v^3}, \quad (11.4.6)$$

and

$$\nabla_{\mathbf{v}} \nabla_{\mathbf{v}} V = \frac{1}{v} \left(\overset{\leftrightarrow}{\mathbf{1}} - \frac{\mathbf{v}\mathbf{v}}{v^2} \right), \quad (11.4.7)$$

to rewrite equation (11.4.4) in the form

$$\frac{\delta_c f_e}{\delta t} = n_0 \Gamma_{ei} \left[-\nabla_{\mathbf{v}} \cdot \left(-\frac{f_e \mathbf{v}}{v^3} \right) + \frac{1}{2} \nabla_{\mathbf{v}} \nabla_{\mathbf{v}} : \left(f_e \frac{v^2 \overset{\leftrightarrow}{\mathbf{1}} - \mathbf{v}\mathbf{v}}{v^3} \right) \right]. \quad (11.4.8)$$

If we now use the identity

$$\nabla_{\mathbf{v}} \cdot \left(\frac{v^2 \overset{\leftrightarrow}{\mathbf{1}} - \mathbf{v}\mathbf{v}}{v^3} \right) = -\frac{2\mathbf{v}}{v^3}, \quad (11.4.9)$$

it is easy to see that the first term in the parentheses, which represents the dynamical friction vector, is cancelled exactly by an identical term of opposite sign from the diffusion tensor. This cancellation is a consequence of the assumption of infinite mass of the ions, which in turn implies that the electron energy is conserved in a collision with ions, allowing no energy transfer to ions. As a result of this cancellation, we are left with the expression

$$\frac{\delta_c f_e}{\delta t} = \frac{n_0 \Gamma_{ei}}{2} \nabla_{\mathbf{v}} \cdot \left(\frac{v^2 \overset{\leftrightarrow}{\mathbf{1}} - \mathbf{v}\mathbf{v}}{v^3} \cdot \nabla_{\mathbf{v}} f_e \right). \quad (11.4.10)$$

We note that the right-hand side of equation (11.4.10) vanishes identically for an isotropic electron distribution function, that is, when $f_e = f_e(v^2)$. This can be seen easily by substituting the identity

$$\nabla_{\mathbf{v}} f_e(v^2) = 2\mathbf{v} \frac{\partial f_e}{\partial v^2}, \quad (11.4.11)$$

into equation (11.4.10), whereby we obtain

$$(v^2 \overset{\leftrightarrow}{\mathbf{1}} - \mathbf{v}\mathbf{v}) \cdot \mathbf{v} = v^2 \mathbf{v} - \mathbf{v} v^2 = 0, \quad (11.4.12)$$

thereby annihilating the collision operator. The physical significance of this result is that the collision operator (11.4.10) will tend to relax an anisotropic electron distribution to an isotropic distribution as collisional processes evolve the plasma toward a steady state.

We now proceed with the calculation of the plasma conductivity [see Goldston and Rutherford, 1995]. We assume that a small electric field $\mathbf{E} = E\hat{\mathbf{z}}$ is applied to a homogeneous currentless plasma in which the electron distribution function is initially Maxwellian (before the electric field is applied). After the transients have died down, let us assume that the plasma evolves to a spatially homogeneous steady state. From equations (11.3.7) and (11.4.10), we obtain

$$-\frac{e}{m_e} \mathbf{E} \cdot \frac{\partial f_e}{\partial \mathbf{v}} = \frac{n_0 \Gamma_{ei}}{2} \nabla_{\mathbf{v}} \cdot \left(\frac{v^2 \overset{\leftrightarrow}{\mathbf{1}} - \mathbf{v}\mathbf{v}}{v^3} \cdot \nabla_{\mathbf{v}} f_e \right). \quad (11.4.13)$$

It is convenient to write the above equation in spherical polar velocity coordinates (v, θ, ϕ) , where $v_x = v \sin \theta \cos \phi$, $v_y = v \sin \theta \sin \phi$, and $v_z = v \cos \theta$. Since the applied electric field points in the z direction and there is no other preferred direction, we require that the distribution function be azimuthally symmetric, that is, $\partial f_e / \partial \phi = 0$. Under these conditions, equation (11.4.13) can be written in the form

$$-\frac{e}{m_e} E \left(\cos \theta \frac{\partial f_e}{\partial v} - \frac{\sin \theta}{v} \frac{\partial f_e}{\partial \theta} \right) = \frac{n_0 \Gamma_{ei}}{2v^3 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f_e}{\partial \theta} \right). \quad (11.4.14)$$

For a small electric field E , we can linearize equation (11.4.14) by writing

$$f_e(v, \theta) = f_{e0}(v) + f_{e1}(v, \theta), \quad (11.4.15)$$

and neglecting terms proportional to the product $E f_{e1}$ on the left-hand side. We assume that the zero-order distribution function is a Maxwellian

$$f_{e0}(v) = n_0 \left(\frac{m_e}{2\pi\kappa T_e} \right)^{3/2} \exp\left(-\frac{m_e v^2}{2\kappa T_e}\right). \quad (11.4.16)$$

The linearized version of equation (11.4.14) is

$$-\frac{e}{m_e} E \cos \theta \frac{\partial f_{e0}}{\partial v} = \frac{n_0 \Gamma_{ei}}{2v^3 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f_{e1}}{\partial \theta} \right). \quad (11.4.17)$$

Substituting equation (11.4.16) into the left-hand side of the above equation, we obtain

$$\frac{eEv \cos \theta}{\kappa T_e} f_{e0} = \frac{n_0 \Gamma_{ei}}{2v^3 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial f_{e1}}{\partial \theta} \right). \quad (11.4.18)$$

Equation (11.4.18) can be solved for $f_{e1}(v, \theta)$ by assuming that

$$f_{e1}(v, \theta) = A v^4 f_{e0}(v) \cos \theta, \quad (11.4.19)$$

where A is a constant to be determined by direct substitution in equation (11.4.18). After carrying out the necessary calculation it is easy to show that

$$A = -\frac{eE}{n_0 \Gamma_{ei} \kappa T_e}. \quad (11.4.20)$$

From the expression for $f_{e1}(v, \theta)$ we can calculate the linearized current density along the direction of the electric field:

$$\begin{aligned} J &= -e \int_{-\infty}^{\infty} f_{e1} v \cos \theta d^3 v \\ &= \frac{e^2 E}{n_0 \Gamma_{ei} \kappa T_e} \int_0^{2\pi} d\phi \int_0^\pi \sin \theta \cos^2 \theta d\theta \int_0^\infty v^7 f_{e0} dv. \end{aligned} \quad (11.4.21)$$

Using equations (11.4.5) and (11.4.16) and carrying out the integrals in the above equation, we obtain

$$J = \frac{32\pi^{1/2} \epsilon_0^2 (2\kappa T_e)^{3/2}}{m_e^{1/2} e^2 \ln \Lambda} E. \quad (11.4.22)$$

The plasma conductivity, σ , defined by the relation

$$J = \sigma E, \quad (11.4.23)$$

Table 11.1 Comparison of the resistivity of various types of plasmas with some common materials

Material	Resistivity, η (ohm m)
Copper	2×10^{-8}
Stainless steel	7×10^{-7}
100 eV plasma	5×10^{-7}
5 keV plasma	1×10^{-9}
Interstellar gas (1 eV)	5×10^{-7}
Solar corona (10 eV)	5×10^{-5}
Earth's ionosphere (0.1 eV)	2×10^{-2}

is then given, in the Lorentz gas approximation, by the expression

$$\sigma = \frac{32\pi^{1/2}\epsilon_0^2(2\kappa T_e)^{3/2}}{m_e^{1/2}e^3 \ln \Lambda}. \quad (11.4.24)$$

It turns out that equation (11.4.21), which is derived assuming that electron–electron collisions are neglected, overestimates the actual plasma conductivity. This occurs not because electron–electron collisions contribute additively to the effective number of collisions that electrons suffer, but more subtly because electron–electron collisions change the electrons distribution function and hence, the electron–ion collisional interaction. Note that, except for the weak dependence on $\ln \Lambda$, the conductivity depends only on the electron temperature, and is completely independent of the number density.

The plasma resistivity η is the inverse of the conductivity σ . A more detailed calculation of the resistivity that takes into account the effect of electron–electron collisions, first carried out by Spitzer and Harm [1953], yields the following formula for the resistivity,

$$\eta \approx 5.2 \times 10^{-5} \frac{\ln \Lambda}{(\kappa T_e)^{3/2}} \text{ ohm m}, \quad (11.4.25)$$

where κT_e is expressed in electron volts. To be precise, the above equation gives the Spitzer–Harm resistivity parallel to an equilibrium magnetic field. The resistivity perpendicular to the magnetic field is approximately twice the parallel resistivity.

It is interesting to compare the resistivity of fully ionized plasmas at different temperatures to some common materials. For example, in Table 11.1 we compare the resistivities of 100 eV and 5 keV plasmas to those of copper and stainless steel. As one can see, a 100 eV plasma has very low resistivity, comparable to that of stainless steel, but larger than that of copper. However, a 5 keV plasma, typical of large laboratory fusion experiments, is an order of magnitude less resistive than copper.

From equation (??), we can make an estimate of the electron–ion collision frequency, ν_{ei} , given by

$$\nu_{ei} \simeq \frac{n_0 e^2}{m_e \sigma} \simeq \frac{n_0 e^4 \ln \Lambda}{32\pi^{1/2} (2kT_e)^{3/2}}. \quad (11.4.26)$$

This equation was previously used in Section ??.

11.5 Collision operator for Maxwellian distributions of electrons and ions

In Section 11.4, we calculated the Rosenbluth potentials in the Lorentz gas approximation when the ion distribution function can be approximated by a Dirac delta function, and ions are assumed to be infinitely massive compared with electrons ($m_e/m_i \rightarrow 0$). In this approximation, there is no energy transfer to ions. We now generalize this treatment to a case when the electrons and ions both have Maxwellian distributions (with different temperatures) and the ions have a large but finite mass. For both the electrons and ions we assume that the distribution function is given by

$$f_s(\mathbf{v}) = n_0 \left(\frac{m_s}{2\pi\kappa T_s} \right)^{3/2} \exp\left(-\frac{m_s v^2}{2\kappa T_s}\right). \quad (11.5.1)$$

Using the identity

$$\int_{-\infty}^{\infty} \frac{\exp(-x^2)}{|\mathbf{x} - \mathbf{y}|} d^3x = \frac{\pi^{3/2}}{y} \operatorname{erf}(y), \quad (11.5.2)$$

where $\operatorname{erf}(y)$ is the error function, defined by

$$\operatorname{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y \exp(-x^2) dx, \quad (11.5.3)$$

we obtain the first Rosenbluth potential

$$H_s(\mathbf{v}) = n_0 \left(\frac{m_s}{2\kappa T_s} \right)^{1/2} \frac{\operatorname{erf}(y)}{y}, \quad (11.5.4)$$

where $y = v/(2\kappa T_s/m_s)^{1/2}$. Similarly, the second Rosenbluth potential can be shown to be

$$G_s(\mathbf{v}) = \frac{n_0}{2y} \left(\frac{2\kappa T_s}{m_s} \right)^{1/2} \left[y \frac{d}{dy} \operatorname{erf}(y) + (1 + 2y^2) \operatorname{erf}(y) \right]. \quad (11.5.5)$$

The dynamical friction vector, given by equation (11.3.20), can now be written

$$\left\langle \frac{\Delta \mathbf{v}}{\Delta t} \right\rangle_s = \sum_{s'} \mathbf{F}_{ss'}, \quad (11.5.6)$$

where the dynamical friction force acting on species s due to species s' is given by

$$\mathbf{F}_{ss'} = -\Gamma_{ss'} n_0 \left[\operatorname{erf}(y) - y \frac{d}{dy} \operatorname{erf}(y) \right] \frac{\mathbf{v}}{v^3}. \quad (11.5.7)$$

Similarly, after some algebra, the diffusion tensor can be shown to be

$$\left\langle \frac{\Delta \mathbf{v} \Delta \mathbf{v}}{\Delta t} \right\rangle_s = \sum_{ss'} \Gamma_{ss'} \frac{n_0}{2y^3} \left(\frac{m_{s'}}{2\kappa T_{s'}} \right)^{1/2} \frac{\mu_{ss'}}{m_s} \left[\frac{\leftrightarrow}{2} g_1(y) + 3 \frac{\mathbf{v} \mathbf{v}}{v^2} g_2(y) \right], \quad (11.5.8)$$

where

$$g_1(y) = (2y^2 - 1) \operatorname{erf}(y) + y \frac{d}{dy} \operatorname{erf}(y), \quad (11.5.9)$$

and

$$g_2(y) = \left(1 - \frac{2y^2}{3} \right) \operatorname{erf}(y) - y \frac{d}{dy} \operatorname{erf}(y). \quad (11.5.10)$$

The magnitude of the dynamical friction force $F_{ss'}$ given by equation (11.5.7) is plotted as a function of the dimensionless velocity parameter y in Figure 11.2. This plot shows that the magnitude of the drag force initially increases linearly with velocity. As the velocity approaches the thermal speed (which corresponds to $y = 1$), the drag force reaches a maximum and then starts to decrease, eventually varying as $1/v^2$ at velocities much greater than the thermal velocity. The occurrence of a well-defined maximum in the drag force leads to the phenomenon of runaways, as described below.

When an electric field \mathbf{E} is applied to a plasma, the electrons and ions acquire different flow velocities or drifts parallel to \mathbf{E} . Assuming that the plasma is spatially uniform, the momentum equation for the electrons can be written

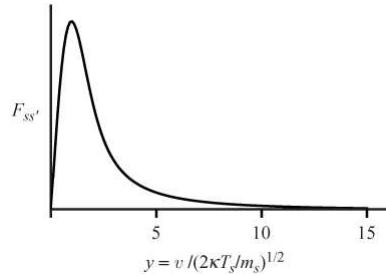


Fig. 11.2 A plot of the dynamical friction force as a function of the dimensionless velocity parameter y .

$$m_e n_0 \frac{\partial \mathbf{U}_e}{\partial t} = -en_0 \mathbf{E} + m_e n_0 \mathbf{F}_{ei}, \quad (11.5.11)$$

where $m_e n_0 \mathbf{F}_{ei}$ is the drag force exerted on the electrons by the ions and \mathbf{F}_{ei} is given by equation (11.5.7). For small electric fields, the drag force increases linearly with v . In this

regime it is possible to attain a steady state whereby the left-hand side vanishes and the applied electric field is exactly balanced by the drag force. Indeed, it is this force balance that produces the condition $\mathbf{J} = \sigma\mathbf{E}$ used in Section 11.4 to define the plasma conductivity. However, if the electric field is larger than a critical field, E_c , given by

$$E > E_c = m_e(F_{ei})_{\max}/e \quad (11.5.12)$$

where $(F_{ei})_{\max}$ is the maximum of the curve in Figure 11.2 (applied to the electrons and ions), then the magnitude of U_e will grow without bound. The critical field E_c is called the Dreicer field [Dreicer, 1959]. For electric field strengths greater than the Dreicer field, electrons are accelerated to sufficiently large speeds that the magnitude of the drag force starts to decrease as $1/v^2$, making it impossible for the drag force to balance the electric field and produce a steady state. Under such conditions, the electron distribution function accelerates without limit. Such runaway acceleration is believed to be responsible for the high velocity electron beams that are frequently observed in both space and laboratory plasmas.

Problems

- 11.1. Calculate numerical values of $\ln \Lambda$ for hydrogen plasmas in the range of density $1 - 10^{20} \text{ cm}^{-3}$ and temperature $10^2 - 10^8 \text{ K}$. How sensitive is $\ln \Lambda$ to such a wide range of density and temperature?
- 11.2. Show that the first and second Rosenbluth potentials, given by

$$H_s(\mathbf{v}) = \int_{-\infty}^{\infty} \frac{f_s(\mathbf{v}')}{|\mathbf{v} - \mathbf{v}'|} d^3v', \text{ and}$$

$$G_s(\mathbf{v}) = \int_{-\infty}^{\infty} f_s(\mathbf{v}') |\mathbf{v} - \mathbf{v}'| d^3v',$$

are related by

$$H_s(\mathbf{v}) = \frac{1}{2} \nabla_{\mathbf{v}} \cdot \nabla_{\mathbf{v}} G_s(\mathbf{v}).$$

- 11.3. If the velocity distribution function for the s th species is a Maxwellian

$$f_s(\mathbf{v}) = n_0 \left(\frac{m_s}{2\pi k T_s} \right)^{3/2} \exp \left(-\frac{m_s v^2}{2k T_s} \right),$$

show that the first Rosenbluth potential is given by

$$H_s(\mathbf{v}) = n_0 \left(\frac{m_s}{2k T_s} \right)^{3/2} \frac{\operatorname{erf}(y)}{y}$$

where $\operatorname{erf}(y)$ is the error function, and $y = v/(2k T_s/m_s)^{1/2}$, and that the second Rosenbluth potential is given by

$$G_s(\mathbf{v}) = \frac{n_0}{2y} \left(\frac{2\kappa T_s}{m_s} \right)^{3/2} \left[y \frac{d}{dy} \operatorname{erf}(y) + (1 + 2y^2) \operatorname{erf}(y) \right].$$

11.4. For the Maxwellian velocity distribution in problem 11.3, show that the diffusion tensor is given by

$$\left\langle \frac{\Delta \mathbf{v} \Delta \mathbf{v}}{\delta t} \right\rangle_s = \sum_s \Gamma_{ss'} \frac{n_0}{2y^3} \left(\frac{m_{s'}}{2\kappa T_{s'}} \right)^{1/2} \frac{\mu_{ss'}}{m_s} \left[\overset{\leftrightarrow}{\mathbf{1}} g_1(y) + 3 \frac{\mathbf{v}\mathbf{v}}{v^2} g_2(y) \right],$$

where

$$g_1(y) = (2y^2 - 1) \operatorname{erf}(y) + y \frac{d}{dy} \operatorname{erf}(y),$$

and

$$g^2(y) = \left(1 - \frac{2y^2}{3} \right) \operatorname{erf}(y) - y \frac{d}{dy} \operatorname{erf}(y).$$

References

- Dreicer, H. 1959. Electron and ion runaway in a fully ionized gas. *Phys. Rev.* **115**, 242–249.
- Goldston, R. J., and Rutherford, P. H. 1995. *Introduction to Plasma Physics*. Institute of Physics: Bristol.
- Ichimaru, S. 1994. *Statistical Plasma Physics*, Vol. II. Addison-Wesley: Reading, MA.
- Rosenbluth, M. N., MacDonald, W. M., and Judd, D. 1957. Fokker–Planck equation for an inverse-square force. *Phys. Rev.* **107**, 1–6.
- Spitzer, L., and Harm, R. 1953. Transport phenomena in a completely ionized gas. *Phys. Rev.* **89**, 977–981.

Further reading

- Boyd, T. J. M., and Sanderson, J. J. *The Physics of Plasmas*. Cambridge University Press: Cambridge, 2003, Chapter 8.
- Schmidt, G. *Physics of High Temperature Plasmas*. Academic Press: New York, 1979, Chapter 11.
- Sturrock, P. A. *Plasma Physics*. Cambridge University Press: Cambridge, 1994, Chapter 10.

Appendix A Symbols

The following list gives all of the mathematical symbols used in this book and a simple description of the quantity involved. The symbols are organized in alphabetical order, with English symbols first and Greek symbols second. In some cases the same symbol has more than one meaning. For example, the symbol n can mean number density, and index of refraction, as well as the integer n . Usually the meaning is obvious from the context. However, to aid in determining the correct meaning the nearest equation and page numbers in which the symbol first appears are given.

Symbol	Definition	Equation	Page
a_0	parameter that controls the width of Gaussian envelope	4.1.3	76
A	area of current loop	3.1.8	25
A	term in equation for index of refraction	4.4.16	97
A	anisotropy factor	9.3.59	381
$A(\mathbf{r} - \mathbf{v}_g t)$	amplitude of wave packet	4.1.25	81
\mathbf{A}	vector potential	3.7.2	61
b	impact parameter	3.7.27	67
b_0	impact parameter for 90° scattering	11.1.2	416
B	magnetic field magnitude	2.4.1	12
B	term in equation for index of refraction	4.4.17	97
B_m	magnetic field magnitude at mirror point	3.4.19	43
B_{\max}	maximum magnetic field strength	3.4.23	45
\mathbf{B}	magnetic field vector	3.1.1	24
\mathbf{B}_0	externally imposed magnetic field	4.4.1	94
$\hat{\mathbf{B}}$	unit vector in direction of magnetic field	3.3.7	33
c	speed of light	4.1.15	79
c_V	specific heat capacity per unit volume	7.3.56	273
C	constant in energy equation	6.7.36	232
C	integration contour in complex v_z space	8.2.9	293
C_A	ion acoustic speed	8.5.31	333
C_s	thermal speed of s th species	2.1.3	6
C_P	heat capacity at constant pressure	6.1.25	181
C_V	heat capacity at constant volume	6.1.25	181
$d\Omega$	differential solid angle	2.5.2	13
ds	distance along magnetic field line	4.4.33	103

d	scale length in Harris current sheet	6.8.1	240
D	dielectric tensor element	4.4.8	95
$D_q(v_z, t)$	quasi-linear diffusion coefficient	10.1.44	399
D_W	whistler dispersion	4.4.34	104
$D(\mathbf{ik}, -i\omega)$	dispersion relation	4.1.13	78
$D_\ell(k, \omega)$	dispersion relation for the longitudinal mode	5.5.18	158
\mathbf{D}	displacement field	4.2.1	84
\bar{D}_a	ambipolar diffusion coefficient	5.7.12	168
\bar{D}_s	diffusion coefficient of s th species	5.7.3	167
$\overset{\leftrightarrow}{D}_s$	matrix representation of a system of linear equations	4.1.19	79
e	electronic charge	2.2.1	7
e_s	charge for s th species	2.3.6	11
E	electric field magnitude	2.3.1	10
$\mathcal{E}(k, t)$	electric field spectral density	10.1.42	398
\mathbf{E}	electric field vector	3.2.2	26
f_{ce}	electron cyclotron frequency, Hz	2.4.2	12
f_{pe}	electron plasma frequency, Hz	2.3.5	11
$f_s(\mathbf{v})$	velocity distribution function for the s th species	2.1.1	5
F	term in the equation for the index of refraction	4.4.18	97
$F_0(v_z)$	normalized one-dimensional velocity distribution function	8.1.12	283
$g(\mathbf{r}, \mathbf{v})$	weighting factor in phase space	5.1.3	138
$g(z)$	complex function	8.2.9	293
$G(z)$	complex function, $G(z) = g(z)(z - z_0)^n$	8.2.9	293
$G_s(\mathbf{v})$	second Rosenbluth potential	11.3.33	425
$G_{s0}(v_{ })$	reduced one-dimensional velocity distribution function	9.3.34	373
h	enthalpy	6.4.13	194
\hbar	Planck's constant divided by 2π	2.7.1	17
H	Hamiltonian	3.7.1	61
$H(\omega/k)$	function used to evaluate beam instabilities	8.1.30	286
$H_s(\mathbf{v})$	first Rosenbluth potential	11.3.18	423
H_S	scale height	6.7.53	236
\mathcal{H}	magnetic helicity	6.6.20	209
\mathbf{H}	magnetic intensity	4.2.1	84
I	current	3.1.7	25
I	integral invariant	3.6.19	52
$I_n(x)$	modified Bessel function of order n	9.2.28	351
J	action integral	3.6.1	48
J	2nd adiabatic invariant	3.6.16	51
J	Jacobian	5.2.5	142
$J_n(x)$	Bessel function of order n	9.2.11	347
\mathbf{J}	current density	3.2.10	28

\mathbf{J}_G	gradient drift current	3.3.13	38
\mathbf{J}_m	magnetization current	3.3.12	38
\mathbf{J}_r	real current	4.2.1	84
k	wave number	4.1.1	75
k_0	wave number at spectral peak	4.1.3	76
k_{\parallel}	parallel component of wave vector	3.6.26	55
k_{\perp}	perpendicular component of wave vector	9.1.12	344
\mathbf{k}	wave vector, also called propagation vector	4.1.8	77
K	kinetic energy	6.4.17	194
$\hat{\mathbf{K}}$	dielectric tensor	4.2.6	86
L	Lagrangian	3.7.1	61
L	dielectric tensor element	4.4.11	96
m_s	atomic mass for the s th species	2.1.1	5
M	magnetic moment	3.7.22	66
M_A	Alfvén Mach number	7.3.16	262
M_S	sonic Mach number	7.3.17	262
\mathbf{M}	magnetization (magnetic moment per unit volume)	3.1.10	26
n	index of refraction	4.1.28	83
n	number of roots of $D(p) = 0$	8.5.12	323
n_A	Alfvén index of refraction	4.4.28	100
n_L	index of refraction for left-hand polarized mode	4.4.43	108
n_n	number density of neutral gas	2.5.1	13
n_R	index of refraction for right-hand polarized mode	4.4.42	108
n_s	number density for the s th species	2.1.2	5
\mathbf{n}	index of refraction vector	4.2.9	86
$\hat{\mathbf{n}}$	unit normal vector	7.1.6	253
N	total number of particles in phase space	5.1.2	138
N_c	average number of collisions per unit path length	11.2.1	418
N_D	number of electrons per Debye cube	2.5.5	15
N_{ℓ}	line number density	6.6.40	213
N_{ℓ}	number of turns per unit length	3.3.10	36
N_w	winding number	8.5.10	322
P	complex frequency ($p = \gamma - i\omega$)	8.2.1	291
p	momentum	3.6.4	48
p_i	conjugate momentum coordinate	3.7.1	61
P	dielectric tensor element	4.4.9	95
$P \int$	principal value integral	8.2.39	301
$P(x)$	integrating factor	9.2.2	345
P_{ij}	pressure tensor	5.1.14	139
$P_s(\mathbf{v}, \Delta\mathbf{v})$	probability distribution function	11.3.1	420
$\hat{\mathbf{P}}_s$	pressure tensor	5.1.7	138
q	charge	2.2.2	7
q	position coordinate	3.6.3	48

q_i	conjugate position coordinate	3.7.1	61
$q(\rho)$	q -factor in screw pinch	6.6.52	215
Q	test charge	2.2.9	8
$Q(x)$	source term in a linear differential equation	9.2.2	345
\mathbf{Q}_s	kinetic energy flux	5.4.31	151
r	radius (in spherical coordinates)	2.2.7	8
R	dielectric tensor element	4.4.11	96
$\text{Res}[g(z), z_0]$	residue of $g(z)$ evaluated at z_0	8.2.9	293
R_m	magnetic Reynolds number	6.3.5	185
\mathbf{R}_C	radius of curvature vector	3.3.8	34
$\hat{\mathbf{R}}_C$	unit vector in direction of instantaneous radius of curvature	3.3.9	34
s	species	2.1.1	5
S	dielectric tensor element	4.4.8	95
S	Lundquist number	6.8.5	241
\mathbf{S}	Poynting flux ($\mathbf{S} = (1/\mu_0) \mathbf{E} \times \mathbf{B}$)	6.4.11	193
t	time	2.3.2	11
T_{rs}	magnetic pressure tensor element	6.2.3	183
T_s	temperature for the s th species	2.1.1	5
T_{\parallel}	parallel temperature	5.1.16	140
T_{\perp}	perpendicular temperature	5.1.16	140
$\overset{\leftrightarrow}{\mathbf{T}}$	magnetic pressure tensor	6.2.4	183
\mathbf{U}_s	average velocity of s th species	5.1.5	138
v	velocity magnitude	2.1.1	5
v_x	x component of velocity	3.4.4	40
v_y	y component of velocity	3.4.4	40
v_z	z component of velocity	3.4.4	40
v_{\parallel}	velocity component parallel to the magnetic field	3.2.1	26
v_{\perp}	velocity perpendicular to the magnetic field	3.1.3	24
$v_{\parallel\text{Res}}$	cyclotron resonance velocity	9.3.39	374
\mathbf{v}	velocity	2.1.1	5
\mathbf{v}_C	curvature drift velocity	3.3.8	34
\mathbf{v}_E	$\mathbf{E} \times \mathbf{B}$ drift velocity	3.2.3	27
\mathbf{v}_F	drift due to a perpendicular external force	3.2.11	31
\mathbf{v}_g	group velocity	4.1.25	81
\mathbf{v}_G	gradient drift velocity	3.3.7	33
\mathbf{v}_p	phase velocity	4.1.9	77
V	volume	5.1.4	138
V	equilibrium potential of the surface	2.2.12	10
V	velocity space	5.4.1	146
V_0	velocity in ring distribution	9.2.32	355
V_j	beam velocity	8.1.26	285
V_A	Alfvén speed	6.5.14	196

V_S	speed of sound	6.5.5	195
ω_{\parallel}	parallel kinetic energy of a particle	3.3.9	34
ω_{\perp}	perpendicular kinetic energy of a particle	3.1.9	25
W	total kinetic energy density	5.4.35	151
W	potential energy density	6.4.18	194
W_s	kinetic energy density	5.4.29	150
W_{\perp}	perpendicular kinetic energy density	3.3.11	38
\mathbf{W}_s	diffusion velocity for species s	6.1.7	177
x	x component of a position vector	3.4.3	40
x	real part of ζ	8.3.7	309
X	dimensionless parameter in CMA diagram	4.4.58	122
X_{WKB}	WKB solution of the time independent harmonic oscillator equation	3.6.7	49
y	y component of a position vector	3.4.3	40
y	complex part of ζ	8.3.7	309
Y	dimensionless parameter in CMA diagram	4.4.58	122
z	z component of a position vector	3.4.1	40
z	normalized z component of velocity	8.3.2	308
$Z(\zeta)$	plasma dispersion function	8.3.5	309
$\overset{\leftrightarrow}{\mathbf{1}}$	unit tensor	4.2.6	86
α	pitch angle	3.4.22	44
α	group velocity angle	4.1.29	84
α	scalar function used in force-free equilibria	6.6.9	206
α_0	loss-cone angle	3.4.23	45
α_s	symbol used for $(k_{\parallel}v_{\parallel} - \omega)/\omega_{cs}$	9.1.14	345
β_c	thermal cyclotron radius times k_{\perp}	9.2.31	352
β_s	symbol used for $k_{\perp}v_{\perp}/\omega_{cs}$	9.1.14	345
γ	power law index, also called the polytrope index	5.4.43	154
γ	ratio of heat capacity at constant pressure to heat capacity at constant volume	7.3.56	273
γ	growth rate	8.1.31	287
γ_0	damping constant for the harmonic oscillator	8.2.5	292
Γ_{ei}	defined parameter in the Fokker-Planck collision operator	11.4.5	427
Γ_n	defined function	9.2.30	351
$\Gamma_{ss'}$	constant factor in dynamical friction vector	11.3.17	423
$\langle \Delta p \rangle$	uncertainty in momentum	2.7.1	17
Δs	change in specific entropy	7.3.56	273
Δx	small displacement in the x direction	2.3.1	10
$\langle \Delta x \rangle$	uncertainty in position	2.7.1	17
δ	shock strength parameter	7.3.15	262
$\delta(x)$	Dirac delta function	5.1.12	139
δ_c	collision operator	5.2.14	143

δ_{rs}	Kronecker delta	6.2.3	183
ϵ	depth of potential well	10.1.54	401
ϵ	small quantity	8.2.39	301
ϵ_0	permittivity of free space	2.2.1	7
ζ	dimensionless complex frequency	8.3.2	308
η	resistivity	6.1.19	179
θ	polar angle (in spherical coordinates)	3.7.22	66
θ	wave normal angle	4.1.28	83
θ_{Res}	resonance cone angle	4.4.53	116
κ	Boltzmann's constant	2.1.1	5
λ	wavelength ($2\pi/k$)	4.1.2	75
λ_D	Debye length	2.2.8	8
Λ	plasma parameter	11.2.9	419
μ	magnetic moment	3.1.6	25
μ_0	permeability of free space	3.3.10	36
$\langle \mu_s \rangle$	average magnetic moment for s th species	3.1.10	26
$\bar{\mu}_s$	mobility coefficient for s th species	5.7.3	167
ν_{ei}	electron-ion collision frequency	2.5.5	15
ν_{ns}	neutral collision frequency	2.5.1	13
ν_s	effective collision frequency for s th species	5.6.1	162
ξ	internal energy per unit volume	6.4.12	194
ξ	fluid displacement vector	6.7.12	227
ρ	radial distance from the symmetry axis	3.3.10	36
	in cylindrical coordinates		
ρ_c	cyclotron radius	3.1.4	24
ρ_p	polarization charge	4.2.1	84
ρ_q	charge density	2.2.1	7
ρ_{qs}	charge density of s th species	10.2.7	408
ρ_r	real charge	4.2.1	84
σ	conductivity	5.6.7	164
σ_n	neutral particle collision cross-section	2.5.1	13
σ_s	conductivity of s th species	5.6.8	164
σ_C	Coulomb collision cross-section	2.5.2	13
σ_H	Hall conductivity	5.6.12	165
σ_T	total scattering cross-section	2.5.4	14
$\sigma_{ }$	parallel conductivity	5.6.12	165
σ_{\perp}	perpendicular conductivity	5.6.12	165
$\overset{\leftrightarrow}{\sigma}$	conductivity tensor	4.2.3	85
τ_A	Alfvén time scale	6.8.3	241
τ_R	resistive diffusion time scale	6.8.4	241
ϕ	azimuthal angle (in cylindrical coordinates)	3.4.10	41
ϕ	eikonal	4.5.2	127
Φ	electrostatic potential	2.2.1	7

Φ_B	magnetic flux	3.5.8	47
χ	scattering angle relative to incident beam	2.5.2	13
ψ	effective potential	3.7.16	63
ψ	Faraday polarization angle	4.4.41	108
ψ	flux function	6.6.56	216
ω	frequency (radian s ⁻¹)	3.6.2	48
ω_b	bounce frequency	8.2.49	305
ω_c	cyclotron frequency	2.4.1	12
ω_{cs}	cyclotron frequency of sth species	2.4.3	12
ω_p	plasma frequency	2.3.8	12
ω_{pe}	electron plasma frequency	2.3.4	11
ω_{pi}	ion plasma frequency	2.3.8	12
ω_{ps}	plasma frequency of sth species	2.3.6	11
ω_x	crossover frequency	4.4.35	105
ω_{LH}	lower hybrid frequency	4.4.50	112
ω_{UH}	upper hybrid frequency	4.4.49	112
$\omega_{L=0}$	L = 0 cutoff frequency	4.4.30	101
$\omega_{R=0}$	R = 0 cutoff frequency	4.4.27	100

Appendix B

Vector differential operators

Rectangular coordinates (x, y, z)

$$\nabla\Phi = \hat{\mathbf{x}} \frac{\partial\Phi}{\partial x} + \hat{\mathbf{y}} \frac{\partial\Phi}{\partial y} + \hat{\mathbf{z}} \frac{\partial\Phi}{\partial z},$$

$$\nabla \cdot \mathbf{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z},$$

$$\nabla \times \mathbf{F} = \hat{\mathbf{x}} \left(\frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) + \hat{\mathbf{y}} \left(\frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) + \hat{\mathbf{z}} \left(\frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \right),$$

$$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2}.$$

Cylindrical coordinates (ρ, ϕ, z)

$$\nabla\Phi = \hat{\rho} \frac{\partial\Phi}{\partial\rho} + \hat{\phi} \frac{1}{\rho} \frac{\partial\Phi}{\partial\phi} + \hat{\mathbf{z}} \frac{\partial\Phi}{\partial z},$$

$$\nabla \cdot \mathbf{F} = \frac{1}{\rho} \frac{\partial}{\partial\rho} (\rho F_\rho) + \frac{1}{\rho} \frac{\partial f_\phi}{\partial\phi} + \frac{\partial F_z}{\partial z},$$

$$\nabla \times \mathbf{F} = \hat{\rho} \left(\frac{1}{\rho} \frac{\partial F_z}{\partial\phi} - \frac{\partial F_\phi}{\partial z} \right) + \hat{\phi} \left(\frac{\partial F_\rho}{\partial z} - \frac{\partial F_z}{\partial\rho} \right) + \hat{\mathbf{z}} \frac{1}{\rho} \left(\frac{\partial}{\partial\rho} (\rho F_\phi) - \frac{\partial F_\rho}{\partial\phi} \right),$$

$$\nabla^2\Phi = \frac{\partial^2\Phi}{\partial\rho^2} + \frac{1}{\rho} \frac{\partial\Phi}{\partial\rho} + \frac{1}{\rho^2} \frac{\partial^2\Phi}{\partial\phi^2} + \frac{\partial^2\Phi}{\partial z^2}.$$

Spherical coordinates (r, θ, ϕ)

$$\nabla\Phi = \hat{\mathbf{r}} \frac{\partial\Phi}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial\Phi}{\partial\theta} + \hat{\phi} \frac{1}{r \sin\theta} \frac{\partial\Phi}{\partial\phi},$$

$$\nabla \cdot \mathbf{F} = \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 F_r) + \frac{1}{r \sin\theta} \frac{\partial}{\partial\theta} (F_\theta \sin\theta) + \frac{1}{r \sin\theta} \frac{\partial F_\phi}{\partial\phi},$$

$$\begin{aligned} \nabla \times \mathbf{F} = & \hat{\mathbf{r}} \frac{1}{r \sin\theta} \left[\frac{\partial}{\partial\theta} (F_\phi \sin\theta) - \frac{\partial F_\theta}{\partial\phi} \right] \\ & + \hat{\theta} \frac{1}{r} \left[\frac{1}{\sin\theta} \frac{\partial F_r}{\partial\phi} - \frac{\partial(r F_\phi)}{\partial r} \right] + \hat{\phi} \frac{1}{r} \left[\frac{\partial(r F_\theta)}{\partial r} - \frac{\partial F_r}{\partial\theta} \right], \end{aligned}$$

$$\nabla^2 \Phi = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r\Phi) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2}.$$

Appendix C Vector calculus identities

$$\nabla \cdot \nabla \Phi = \nabla^2 \Phi$$

$$\nabla \cdot \nabla \times \mathbf{F} = 0$$

$$\nabla \times \nabla \Phi = 0$$

$$\nabla \times (\nabla \times \mathbf{F}) = \nabla(\nabla \cdot \mathbf{F}) - \nabla^2 \mathbf{F}$$

$$\nabla(\Phi \Psi) = (\nabla \Phi) \Psi + \Phi \nabla \Psi$$

$$\nabla(\mathbf{F} \cdot \mathbf{G}) = (\mathbf{F} \cdot \nabla) \mathbf{G} + \mathbf{F} \times (\nabla \times \mathbf{G}) + (\mathbf{G} \cdot \nabla) \mathbf{F} + \mathbf{G} \times (\nabla \times \mathbf{F})$$

$$\nabla \cdot (\Phi \mathbf{F}) = (\nabla \Phi) \cdot \mathbf{F} + \Phi \nabla \cdot \mathbf{F}$$

$$\nabla \cdot (\mathbf{F} \times \mathbf{G}) = (\nabla \times \mathbf{F}) \cdot \mathbf{G} - (\nabla \times \mathbf{G}) \cdot \mathbf{F}$$

$$\nabla \times (\Phi \mathbf{F}) = (\nabla \Phi) \times \mathbf{F} + \Phi \nabla \times \mathbf{F}$$

$$\nabla \times (\mathbf{F} \times \mathbf{G}) = (\nabla \cdot \mathbf{G}) \mathbf{F} - (\nabla \cdot \mathbf{F}) \mathbf{G} + (\mathbf{G} \cdot \nabla) \mathbf{F} - (\mathbf{F} \cdot \nabla) \mathbf{G}$$

$$\int_S \mathbf{F} \cdot d\mathbf{A} = \int_V \nabla \cdot \mathbf{F} d^3 x$$

$$\int_C \mathbf{F} \cdot d\ell = \int_S (\nabla \times \mathbf{F}) \cdot d\mathbf{A}$$

Index

- accessibility 121
 - modes accessible from free space 125
 - modes not accessible from free space 125
- action integral 48, 49
- adiabatic equation of state 154
- adiabatic invariants 48–60
 - first 50–51
 - harmonic oscillator 48–50
 - second 51
 - third 53
- Airy functions 130
- Alfvén
 - index of refraction 100, 134
 - Mach number 262, 278
 - mode 135
 - speed 134, 196, 257
 - time scale 241
- Alfvén waves 100
 - generated by Jupiter's moon Io 203
 - solar wind 202
 - transverse 198–199, 248
- Ampère's law 36, 38, 57, 85, 86, 182
- analytic continuation 293, 296, 297
- analytic function 293
- anisotropy factor 377
 - electron 381, 388, 389
 - ion 385
- anti-shielding 9, 411
- aurora 2
- average
 - distance between neighboring particles 16
 - kinetic energy 6
 - number density 138
 - velocity 138
- average (expectation) value 138
- axially symmetric magnetic field 39, 44, 50
- Bennett pinch 211, 247
- Bennett relation 213
- Bernstein–Green–Kruskal (BGK) electrostatic potentials 406, 409–411
- Bernstein-like modes 358
- Bernstein mode 350–355, 387
- Bessel functions 348, 349, 352, 369
 - modified 351
- Bohm–Gross dispersion relation 159, 284–285, 303
- Boltzmann equation 140–144
- bounce frequency 305
- bounce resonance 58
- bounded integrable orbits 68, 71
- bounded volumes in CMA diagram 121
- bounding surfaces in CMA diagram 121
- bow shock 257, 276
- bump-on-tail instability 328–330, 399–400
- Buneman mode 289–290
- Cauchy residue theorem 292
- Cauchy velocity distribution function 298–300, 313–315
- chaos 404
 - Hamiltonian system 68, 402
 - magnetic field 71
 - single particle orbits 68–71
 - velocity space diffusion 400–406
- characteristic parameters of a plasma 5–18
- characteristic resonance energy
 - ion cyclotron mode 384
 - whistler mode 379
- charge continuity equation 147, 183
- charge density 7
- charge neutrality 1, 87, 182, 392
- charge separation 10
- Chew–Goldberger–Low (CGL) equation of state 154–155
- Clemmow–Mullaly–Allis (CMA) diagram 120–127
 - electrons 123
 - electrons and ions 126
- closure problem 152, 394
- collective motions 2, 3
- collisional drag force 162–163
- collision cross-section
 - charged and neutral particles 13
 - charged particles 14, 416
 - total 417
- collision frequency 13–15, 132
 - charged and neutral particles 13
 - charged particles 13–15, 17, 415–433
 - effective 163, 178
 - electron–ion 15, 431
- collision operator 143
- collisions 1, 140, 144, 415–433
 - fully ionized plasma 415
 - isotropic 163
 - large angle 16

- partially ionized plasma 13, 163
 small angle 13–16, 417–420
 complex *D*-plane 320
 complex frequency 395
 complex *p*-plane 293, 296
 conductivity
 fully ionized gas 427–431
 Hall 165
 Lorentz 247
 Ohm's law 163
 parallel 165
 partially ionized gas 164–166
 Pedersen 164–166
 perpendicular 165
 sth species 164
 conductivity tensor 85, 86, 367
 cold magnetized plasma 95
 cold unmagnetized plasma 88
 hot magnetized plasma 370
 conjugate momenta 48, 61, 62, 129
 conservation of energy 252
 conservation of energy flux 254
 conservation of mass 252
 conservation of mass flux 253
 conservation of momentum 252
 conservation of momentum flux 253
 constancy of the magnetic moment 41–43
 constant- α surface 206
 constants of motion 145
 continuity equation in six-dimensional phase space
 142
 contour integral 292
 convective derivative 150, 178, 189
 co-rotational electric field 29, 30
 cosmic rays 66
 Coulomb
 collisions 416–417
 force 2, 13
 potential 9
 scattering 14
 critical drift velocity 365
 critical electric field 433
 crossover frequency 104, 105, 106
 current-driven electrostatic ion cyclotron waves
 363–366
 current-driven ion acoustic instability 331–334, 339
 cusp mirror geometry 226
 cyclic coordinates 62, 129
 cyclotron
 frequency 12, 24
 harmonics 348
 radius 24
 resonance 55
 resonance energy 380
 resonance interactions 360
 resonance velocity 382
 cyclotron damping 341, 374–377
 electron 103, 104, 376
 ion 106, 376–377
 cyclotron frequency 12, 24
 electron 12
 ion 12
 cylindrical magnetic field geometries 211–215
 cylindrical velocity space coordinates 342, 343
 damping constant 292
 de Broglie wavelength 17
 Debye–Hückel potential 8, 9
 Debye length 7–9, 159, 303
 Debye shielding 7–10
 degrees of freedom 61
 de Hoffmann–Teller frame 258–260
 destructive transition in CMA diagram 121
 diamagnetism 26
 dielectric tensor 85, 86, 367
 cold magnetized plasma 95
 cold unmagnetized plasma 88
 hot magnetized plasma 370
 differential rotation of the Sun 191
 diffusion 166
 ambipolar 166–169
 coefficient 167
 equation 168, 191
 tensor 421, 424–426, 432, 434
 velocity 177
 dipole moment per unit volume 84
 Dirac delta function 77
 discontinuities 251, 255
 contact 255–256
 rotational 255, 256–258
 shock 258
 tangential 256
 dispersion 79
 dispersion relation 78–80, 86
 electromagnetic ion cyclotron waves in a hot
 magnetized plasma 384
 electromagnetic waves in a cold magnetized plasma
 97
 electromagnetic waves in a cold unmagnetized
 plasma 89
 electromagnetic waves in a hot magnetized plasma
 367–386
 electron plasma oscillations 92, 159
 electrostatic waves in a cold beam plasma 285–290
 electrostatic waves in a hot unmagnetized plasma
 283–284, 294–298, 299–303
 electrostatic waves in a hot magnetized plasma
 345–366
 electrostatic waves in a magnetized Maxwellian
 plasma 350–355
 electrostatic waves in an unmagnetized Maxwellian
 plasma 308–311
 free space 79

- ion acoustic waves in a hot unmagnetized plasma 160–162
 Langmuir waves in a hot unmagnetized plasma 158–159, 302, 313, 316
 Langmuir waves in an unmagnetized Maxwellian plasma 308–311
 displacement current 85
 displacement field 84
 distribution function (*see* velocity distribution function) 5, 137–140
 bi-Maxwellian 140, 377–378, 388
 Cauchy 170, 298–300, 320, 321, 335
 counter-streaming Cauchy 327, 328, 338
 double-humped 325
 kappa 336
 loss-cone 359, 382
 Maxwellian 5, 139, 169, 308, 337, 338, 387, 388
 non-equilibrium 6
 non-thermal 6
 reduced 295, 312
 ring-type 355
 solar flare electrons 405
 spatially averaged 392
 spherical shell 334, 357
 Doppler shift 55, 153
 double layer 409
 drag force 149, 432
 Dreicer electric field 433
 drift velocities
 curvature drift 34
 $\mathbf{E} \times \mathbf{B}$ drift 27
 gradient drift 32
 near an x-type neutral line 58
 perpendicular force 31–32
 dynamical friction vector 421, 422–424, 431
 Earth’s magnetosphere 2, 29, 56, 58, 102, 225
 bow shock 276
 ionosphere 239
 magnetopause 256–258
 magnetotail 56
 plasmapause 30
 plasmasphere 29, 30, 225
 Van Allen radiation belt 68
 effective potential 43, 63, 64, 66
 eigenfrequency 231
 eigenmodes 231
 eigenvectors 201
 fast magnetosonic mode 201
 slow magnetosonic mode 201
 transverse Alfvén mode 198
 waves in a cold magnetized plasma 98, 110, 111
 waves in a cold unmagnetized plasma 89
 eikonal 127
 electric dipole antenna 93
 electromagnetic
 ion cyclotron waves 100–101, 376
 waves in a cold magnetized plasma 94–127
 waves in a cold unmagnetized plasma 87–94
 waves in a hot magnetized plasma 367–386
 electron acoustic waves 160
 electron and ion pressure waves 155–162
 electron cyclotron harmonics 354–355
 electron phase-space holes 410
 electron plasma oscillations 10–12, 158, 350
 electron plasma oscillations near Jupiter’s bow shock 278
 electronic charge 7
 electrostatic
 electron cyclotron waves 360
 energy density 398
 instabilities 327–334
 ion cyclotron waves 361, 363
 noise at Jupiter’s bow shock 278
 potential 7
 waves in a cold magnetized plasma 98, 111
 waves in a cold unmagnetized plasma 92
 waves in a hot magnetized plasma 341–386
 waves in a hot unmagnetized plasma 281–334
 elementary plane wave 77, 78
 elliptically polarized wave 108
 energy conservation equation 151
 energy dissipation 273
 energy equation (MHD) 192–194
 energy equation for the *s*th species 151
 energy method of stability analysis 231–235
 enthalpy 194, 252, 256, 261
 entropy 181, 273, 274
 equations of state 152, 181–182
 adiabatic 153–154, 181
 Chew-Goldberger-Low (CGL) 154–155, 182
 cold plasma 152–153
 incompressible 181
 isothermal 181
 power law 181, 182
 error function 431, 434
 evanescent wave 91, 97, 100
 external sources 92–94
 extraordinary mode 111
 Faraday rotation 107–109
 Faraday’s law 45, 53, 86, 182
 Fermi acceleration 59
 field aligned currents 226
 field aligned density irregularities 102
 finite difference equation 401
 finite Larmor radius effects 244
 first law of thermodynamics 153
 first-order linearized Vlasov equation 343
 first-order quantity (definition) 87
 fluid-like motions 161
 flux
 coordinates 188
 function 216

- ropes 208
- surfaces in a toroidal magnetic field 215–219
- Fokker–Planck
 - collision operator 421, 426, 427
 - equation 420–421
- Fourier representation of waves 75–84
- Fourier transforms 75
- free energy sources 358, 362, 386
 - beam 386
 - current 363, 365–366
 - loss-cone 362
- free space L-O mode 122–126
- free space R-X mode 122–126
- frozen field theorem 186–191
- fusion 2–3
- Gardner's theorem 318–319, 321
- gaseous state 1–2
- Gaussian harmonic wave packet 76
- Gauss' law 10, 90, 182
- Gauss' theorem 147
- generalized coordinates 61
- generalized Ohm's law 178–181
- gradient drift current 38
- Grad–Shafranov equation 218
- group velocity 80–84, 91
 - Alfvén mode 199
 - whistler mode 133
- growth rate
 - Buneman mode 290
 - current-driven ion cyclotron waves 365
 - electromagnetic ion cyclotron waves 385
 - electromagnetic waves in a hot magnetized plasma 375
 - electrostatic cold beam modes 285–290
 - electrostatic waves in a hot unmagnetized plasma 290–303
 - ion acoustic waves in a hot unmagnetized plasma 314, 332
 - ion acoustic waves in an unmagnetized Maxwellian plasma 315–318
 - Langmuir waves 303, 311
 - whistler waves 382
- guiding center 24
- gyrotropic condition 343
- Hall current 244
- Hall MHD 245
- Hamiltonian 48, 60–62, 128
- Hamiltonian for a magnetic mirror field 62–66
- Hamilton's equations 60, 61, 128
- harmonic oscillator 10, 221, 292
 - equation 11, 48
 - WKB solution 49
- Harris dispersion relation 345–349, 386, 387
- Heisenberg's uncertainty principle 17
- helium whistler 105, 106
- high density approximation 112, 119
- high-phase velocity approximation 284, 339, 361
- homogeneous equation 79, 86, 367
 - electromagnetic waves in a cold magnetized plasma 96
 - electromagnetic waves in a cold unmagnetized plasma 89
 - electromagnetic waves in a hot magnetized plasma 371
 - electrostatic waves in a hot unmagnetized plasma 157, 283
 - magnetohydrodynamic waves 197
- Hooke's law 10
- Hopf's theorem 207, 215
- H-theorem 6
- hybrid resonances 111, 350
 - ion 111, 112, 113
 - lower 111, 112–113
 - upper 111–112, 353, 355, 360, 387
- ideal gas law 140, 153, 167
- ideal MHD 180
- ideal MHD force operator 231
- index of refraction 99, 389
 - surface 83, 116
 - vector 83
- inhomogeneous plasmas 120
- initial value problem 291
- instabilities 355–366
 - anisotropy-driven 355–363, 377–385
 - cold beam 285–290
 - counter-streaming beam 287
 - current-driven 331, 355, 363
 - interchange 221–227
 - kinetic 327
 - Kruskal–Schwartzschild 223, 236
 - Rayleigh–Taylor 223, 236–239
 - sausage-mode 225
 - whistler-mode 378, 388
- integral invariant 52
- integrating factor 345, 369
- internal energy 194, 252
- internally trapped particles 68
- Io 203
 - Alfvén waves 203
 - controlled radio emissions 203
 - plasma torus 225
- ion acoustic speed 333
- ion acoustic waves 158, 160–162, 312–318, 339
- ion cyclotron
 - damping 366, 376, 377
 - heating 107
 - resonance energy 389
 - resonance velocity 384
 - waves 101
 - whistler 104–107, 376
- ion–ion cutoff frequencies 101
- ion–neutral collisions 163

- ionogram 113, 114
 ionosphere 2, 239
 ionospheric sounders 113–114, 354
 ion-to-electron mass ratio 10
 isoentropic 181
 Jacobian 142
 Joule heating 151
 jump conditions 251
 jump symbol 253, 255
 Jupiter’s magnetosphere 203, 225
 bow shock 277–278
 Io plasma torus 225
 radio emission 203
 kinetic energy density 138, 151, 194
 kinetic energy flux 151
 kinetic theory 137–146
 $L = 0$ cutoff frequency 101
 Lagrangian 61
 Landau damping 160, 162, 290–308, 315
 linear 308
 non-linear 305–308
 physical origin 304–308
 Landau resonance velocity 364
 Langmuir waves 10, 158–160, 284
 Langmuir waves generated by a solar flare 405
 Langmuir waves in the solar wind 329, 330
 Laplace’s equation 169
 Laplace transform 281, 291–294, 335
 first derivative 292
 higher order derivatives 292
 integration contour 291, 297
 inverse 291, 294, 295, 298, 335
 second derivative 292
 large-argument expansion of the plasma dispersion function 309–310, 316, 364
 lattice-like structures 17
 left- and right-hand polarized free space modes 107
 left-hand polarized ordinary (L-O) mode 100, 101, 123, 374
 lightning 2, 101, 225
 ball 2
 bead 225
 linear force operator 227–228, 231, 234, 248
 linearization of the Vlasov equation 342–345
 local thermodynamic equilibrium 153
 longitudinal waves 92, 98, 157
 long-range force 2, 140
 long-wavelength approximation 361
 loop antenna 93
 Lorentz force 24, 41
 Lorentz gas model 162–163, 427
 loss cone 45, 172, 380
 low-density approximation 113
 low-phase velocity approximation 339
 low-temperature, high-phase velocity limit 386
 low-temperature, long wavelength limit 350, 373–374
 L-shell 53
 Lundquist number 241
 macroscopic averages 15
 magnetic
 amplification 190
 beach geometry 106
 bottle 3, 43
 confinement 3
 convection 185–191
 diffusion 191
 energy density 194
 field aligned currents 363
 field strength 12
 flux 63
 flux annihilation 241
 flux invariant 47, 53, 54
 helicity 209
 merging 57–58
 mirror field 39
 mirror force 40–41
 mirror machine 362, 380
 mirror ratio 45–72
 moment 25, 26, 41, 66
 pressure 39, 183–185
 pressure tensor 184
 reconnection 240–244, 245
 Reynolds number 185
 magnetization 26
 magnetization current 26, 38, 84
 magnetohydrodynamics (also *see* MHD) 175–245
 magnetohydrodynamic waves 195–204
 fast magnetosonic mode 197
 slow magnetosonic mode 197
 transverse (shear) Alfvén mode 197, 198, 257
 magnetosonic modes (*see* magnetohydrodynamic waves) 200–201
 magnetosphere 2
 Earth 29, 56, 58, 68, 102, 225
 Jupiter 203, 225
 magnetostatic equilibrium 39
 marginal stability condition 318, 388
 marginal stability line 320
 Markov process 420
 mass continuity equation 176, 182
 Maxwellian electron and ion velocity distribution functions 315–318
 Maxwellian velocity distribution function 6, 350
 Maxwell’s equation 40
 mean-free path 244, 251, 419
 method of successive approximations 285
 MHD (also *see* magnetohydrodynamics) 175
 diffusion time scale 241
 discontinuities 251–278
 equations 182–183
 jump conditions 251–255
 shocks 276–278

- stability 219–239
 MHD equilibria
 force balanced 210–219
 force-free 206–210
 microscopic scale length 251
 mirror point magnetic field 52
 mobility coefficient 167
 moment equations 146–151
 moments of the Boltzmann equation 146–151
 first 147–150
 second 150–151
 zeroth 146–147
 momentum equation 150, 176–178, 182
 multi-component plasma 311
 neutron star 2, 190
 Newton’s law 141
 non-conservative electric field 45–46
 non-destructive transition in CMA diagram 121
 non-integrable system 62
 non-linear effects 391–411
 non-linear evolution and stability of MHD shocks 275
 non-propagating wave 91
 normalized one-dimensional velocity distribution function 283
 normal mode method of MHD stability analysis 228–231
 normal modes 79
 number density 5–6, 138
 number of degrees of freedom 154, 181
 number of electrons per Debye cube 15–17, 144, 415
 Nyquist criterion 320–326
 oblique propagation 116–127
 cold magnetized plasma 116–127
 hot magnetized plasma 385–386
 Ohmic heating 164
 Ohm’s law (also *see* conductivity and resistivity) 163–164, 178, 180, 182
 operator substitutions 78
 ordinary mode 110
 parallel kinetic energy 60
 parallel propagation in a cold magnetized plasma 97–101
 parallel propagation in a hot magnetized plasma 372–385
 parallel resonance velocity 362
 Penrose criterion 321–326, 327, 338
 permeability of free space 36
 permittivity of free space 7
 perpendicular kinetic energy 60
 perpendicular propagation
 cold magnetized plasma 110–116
 hot magnetized plasma 350–355
 perturbed kinetic energy of an MHD equilibrium 232
 perturbed potential energy of an MHD equilibrium 232
 phase mixing 305
 phase space 137
 phase space trajectories 304
 phase-space volume element 137, 141, 144
 phase velocity 77, 91
 ϕ -pinch 214
 photoelectron current 10
 pitch angle 43, 44, 381
 pitch-angle scattering 383
 Planck’s constant 18
 planes of constant phase 77
 plane wave 77
 plasma
 definition 1
 fourth state of matter 1
 recombination 1
 representative parameters 18
 sources of ionization 1
 plasma β 185
 plasma dispersion function 308–311, 336, 364
 large-argument expansion 309
 power series expansions 309
 small-argument expansion 309
 plasma frequency 10–12
 electron 11
 ion 161
 sth species 11
 plasma heating
 via hybrid resonances 115–116
 via ion cyclotron resonances 107
 via Ohmic dissipation 164
 plasma parameter 15, 419
 plasmapause 30
 plasma sheaths 9–10
 plasmasphere 29, 30, 225
 Plemelj relation 301, 309, 321, 357, 361, 374
 Poeverlein construction 129, 130
 Poincaré surface of section 70
 Poisson’s equation 7, 281, 282, 392, 395, 406
 polarization 105
 polarization charge 31, 37, 84
 polarization reversal 104, 106, 107
 poles of a complex function 291, 294
 polytrope index 154, 181
 positive ion sheath 9
 Poynting flux 92, 194, 198
 pressure 39
 pressure tensor 138, 148
 principal value integral 301
 principle of superposition 79
 propagating wave 97
 propagation cutoff 90, 99, 116, 129
 propagation vector 77
 proton whistler 105, 106, 376
 Q resonances 355, 387
 q -factor 215
 quantum effects 17

- quasi-linear
 diffusion coefficient 399, 400
 diffusion equation 393–399, 400, 411
 theory 391–406
- $R = 0$ cutoff frequency 99–100, 107
- radial diffusion 60
- radiation belt 45, 66, 362, 380
- radius of curvature 34
- Rankine–Hugoniot equations 261
- ratio of collective to discrete interactions 17
- ratio of heat capacities 153
- ratio of kinetic energy to potential energy 16–17
- ratio of plasma pressure to magnetic field pressure 185
- Rayleigh–Taylor instability 190
- ray path 82, 127–131
- real charges 84
- real current 84
- reconnection 258
- reduced mass 416
- reduced velocity distribution function 283, 373, 405
- refraction 127
- relationship between dielectric tensor and the conductivity tensor 86
- residue theorem 293, 294, 295, 298
- resistive MHD 182–183
- resistivity 179, 430
- resonance cone 116–117, 119
- resonant particles 304
- reverse-field pinch 209, 210
- reversibility 272–274
- right-hand polarized extraordinary (R-X) mode 100, 106, 123, 374
- ring current 35, 36
- root-mean-square velocity 6
- Rosenbluth potentials 433
 first 424, 431, 434
 second 426, 431, 434
- rotational discontinuity 258
- runaway particles 432
- Rutherford scattering cross-section 14, 417, 422
- scale height 236
- scattering cross-section (see Rutherford) 13
- screw pinch 214–215
- secondary electrons 10
- second law of thermodynamics 272, 273
- self-adjoint operator 229, 248
- separatrix
 magnetic field 56
 phase space 305
- shear Alfvén mode (*see* magnetohydrodynamic waves) 197
- shock
 adiabatic 266, 279
 coordinates 261
 electrostatic 409
 fast 266, 271
- field geometry 259
- foot 276
- intermediate 266, 271
- oblique 271–272
- overshoot 276
- parallel 268–269
- perpendicular 269
- propagation speed 260–266, 267
- ramp 276
- slow 266, 271
- strong 270
- switch-on 269
- transition region 276–278
- wave 59, 251, 255, 258–278
- weak 266
- short-range force 140, 144
- single particle motion 23–71
 dipole magnetic field 66–68
 magnetic mirror field 39–68
 near an x-type neutral line 69
 parabolic magnetic field 69, 70
 perpendicular electric and magnetic fields 26–29
 static uniform magnetic field 23–26
 time varying magnetic field 45–47
- small amplitude waves 75
- small-argument expansion of the plasma dispersion function 309, 310, 316, 364
- Snell’s law 129
- solar corona 58, 208
- solar flare 59, 276, 329, 330, 404–405
- solar wind 56, 202, 256, 257, 258
 interaction with a magnetized planet 257
 magnetic field 202, 246, 329
 pressure 60
 velocity 202, 276
- Solov’ev equilibria 218
- sonic Mach number 262
- sound speed 251, 261
- species 5
- spectral density 398
- Spitzer–Harm resistivity 430
- stability
 electromagnetic waves in a hot magnetized plasma 377–385
 electrostatic waves in a hot unmagnetized plasma 318–334
 magnetohydrodynamic 219–239
- Standard Map 401, 402, 403, 404
- static MHD equilibria 204–219
- static MHD equilibrium conditions 205–206
- stationary non-linear electrostatic potentials 406–411
- stellarator 37
- Stokes’ theorem 46
- strongly coupled plasmas 17, 416
- strong shock limit 270
- supernova 59, 276

- Sweet–Parker model of magnetic reconnection 242–244
 tangent form of the cold plasma dispersion relation 97
 temperature 5–6
 thermal cyclotron radius 352
 thermal excitation of hybrid resonances 114–115
 thermal speed 6, 156
 electron 9
 ion 9
 thermodynamic state variables 273
 threshold drift velocity 366
 time stationary electrostatic potentials 145, 304
 tokamak 37, 216
 toroidal magnetic field 36, 37, 215–219
 total pressure 177
 total scattering cross-section 14
 transverse waves 89–92, 98, 157
 trapped particles in phase space 305
 trapping of resonant electrons 383
 turning points 43
 two-stream instability 287, 410
 type III solar radio bursts 329, 330, 404
 ultraviolet radiation 10
 uncertainty relation of quantum mechanics 77
 uncertainty theorem for Fourier transforms 77
 unit tensor 86
 untrapped particles in phase space 305
 Van Allen radiation belts 68
 variational property of the perturbed MHD potential
 energy 233
 vector calculus identities 443
 vector differential operators 441–442
 vector Helmholtz equation 247
 vector potential 61, 209, 216
 violations of the adiabatic invariants 54–60
 first 55–58
 second 58–59
 third 59–60
 virial theorem 204
 Vlasov approach to the analysis of small amplitude
 waves 281–290
 Vlasov equation 144, 281, 392
 solutions based on constants of the motion 144–146
 spatially averaged 393
 time stationary 406
 waveguide 131
 wavelength 76
 wave normal angle 83
 wave number 76
 waves in a cold plasma 75
 waves in a hot magnetized plasma 341–386
 waves in a hot unmagnetized plasma 281–334
 wave steepening 251, 252, 275
 wave vector 77
 weak-beam approximation 288–289
 weak growth rate approximation 300–303, 357, 358, 361, 385
 weak shock limit 266–267
 Wentzel–Kramer–Brillouin (WKB) equation 49, 129, 130
 whistler dispersion 104
 whistler mode 99, 100, 123, 124, 125, 376
 whistler-mode chorus 383
 whistler-mode emissions in planetary radiation belts 383
 whistler-mode hiss 383
 whistlers 100, 101–107
 winding number 322
 winding theorem 322, 324
 x-type magnetic field configuration 240
 x-type neutral line 56, 58, 69, 242
 x-type neutral point 56, 69
 zero-order quantities 87
 Z-mode 123, 124, 125
 Z-pinch 211–214