https://theses.gla.ac.uk/

Theses Digitisation:

https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses

https://theses.gla.ac.uk/

research-enlighten@glasgow.ac.uk

Propagation of Waves in Inhomogeneous Media

Thesis submitted to the University of Glasgow

by

Carolyn Cunningham Sellar

for the Degree of Doctor of Philosophy

The Department of Physics and Astronomy,

University of Glasgow,

Glasgow G12 8QQ

October 1989

© C.C. Sellar, 1989.

ProQuest Number: 10999283

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

ProQuest 10999283

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.
This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To Mum, Dad and David

Contents

Acknowledgements

I would like to thank Professor I.S. Hughes for the provision of facilities in the Department of Physics and Astronomy of the University of Glasgow.

I am also grateful to Professor E.W. Laing for his role in helping me to complete this thesis and to Professor J.C. Brown for his contributions to the work of Chapter 4. Many thanks are due to Dr. D.A. Diver who made time for frequent discussions and made many helpful suggestions on all aspects of my research, from theory to running the computer.

My parents have provided constant support and encouragement throughout 20 years of full time education. Although it may not always have been apparent, I really appreciate their help and unselfish consideration.

Above all, I would like to thank my fiancé, David Alexander, for the care with which he read this manuscript - I take full responsibility for any errors which remain. He has encouraged me throughout my postgraduate studies and has never complained about the time which I required to complete this thesis. He has listened to my problems, cheered me up and given me lots of practical help. I hope that I can do as much for him.

Abstract

Waves propagating in an inhomogeneous medium differ from waves in uniform surroundings through the dependence of their properties on the variation of the physical parameters of the medium. In this thesis, we will investigate the effects of inhomogeneity on wave propagation in two different cases - introducing the subject via propagation of waves in non-uniform atmospheres and culminating in a full analytic solution of the cold plasma equations describing wave propagation in a plasma with a spatially rotating magnetic field. Plasmas can support a wide variety of waves. In non-uniform plasmas, it is of great interest to consider the possibility of one type of wave undergoing *mode conversion* to a completely different wave - a phenomenon used to heat plasmas in fusion reactor experiments.

In Chapter 1, we present an overview of plasmas in general and consider their widespread natural occurrence and the vast range of their characteristic parameters. This chapter also contains the definition of certain basic plasma quantities, such as *plasma frequency*, which will be used extensively in later chapters and a discussion of plasma confinement systems, illustrating the magnetic field configurations of interest in the context of fusion reactors.

Mode conversion is introduced formally in Chapter 2 where various approaches are discussed. The philosophy behind the powerful WKBJ theory, which applies to slowly-varying media, is presented as a natural extension of the description of waves in uniform media. The "local dispersion relation" method favoured by many authors is examined critically and an alternative description is outlined which is derived from the full differential equation including the gradient terms.

The equations of fluid theory are derived from kinetic theory in Chapter 3 using the method of moments. We discuss how a cold (pressureless) plasma may be described satisfactorily using these equations. Thus, we conclude that the same basic set of equations may be applied to propagation of waves in a cold plasma and a neutral atmosphere (with the removal of electromagnetic forces and addition of a gravitational field).

As a simple introduction to the study of wave propagation in inhomogeneous media, we consider the propagation of waves in an atmosphere where the temperature (and hence sound speed) varies with height. Oscillations in an isothermal atmosphere are shown to possess two branches corresponding to acoustic and gravity

waves. In considering non-uniform atmospheres, the work of authors in the fields of atmospheric and solar physics is combined and reviewed. The isothermal dispersion relation is demonstrated to be inadequate in describing waves in temperature stratified atmospheres. It is found that the ordinary differential equation constructed only has solutions for a limited number of special cases of temperature variation and, in general, requires numerical solution.

Chapter 5 contains the analysis of the propagation of cold plasma waves in a constant magnetic field. In particular, we examine propagation perpendicular to the magnetic field direction in order to provide the background for the equivalent inhomogeneous case of Chapter 6.

In Chapter 6 we solve the problem of wave propagation in a spatially rotating magnetic field. It is shown that, in order to balance the gradient of the equilibrium magnetic field, a current is required which we partition between the ions and electrons. The number of non-vanishing equilibrium quantities is therefore considerably extended from the uniform situation of Chapter 5, leading to a significant alteration of the form of the Ohm's Law which must now contain the electric field plus its first two derivatives. By transforming reference frame, it becomes possible to eliminate the position-dependent coefficients of the differential equations and thus derive a dispersion relation describing waves in such a field structure. It is shown that the waves consist of a propagating part modulated by a periodic envelope induced by the periodicity of the field.

Finally, in Chapter 7, we make suggestions for extensions of the work of Chapter 6.

Chapter 1 - Introduction

1. Plasma - what is it ?

"A plasma is a quasineutral gas of charged and neutral particles which exhibit collective behaviour." But what is the definition of *quasineutral*? And what do we mean by *collective behaviour* ?

It is generally stated that as much as 99% of the matter in the universe is in the plasma state. On the Earth, however, there are so few naturally occurring plasmas because the temperature is seldom high enough to produce the degree of ionisation required for a plasma. Short-lived plasmas may be created in the Earth's atmosphere by the energy released in flashes of lightning but generally we must look further afield for naturally occurring plasmas. In our immediate neighbourhood of the Solar system, for instance, we can observe the Van Allen radiation belts and the solar wind - both composed of highly energetic charged particles. Beyond this, the Sun, the stars and interstellar space contain a wealth of plasmas of widely differing properties.

In the search for new sources of energy, thermonuclear fusion has been proposed as the perfect solution to our needs but there remains a great deal of research to be done and a major role in this will be played by plasmas - the conditions for sustained fusion requiring temperatures and densities not attainable by any other state of matter. A more recent application of plasma physics has been in the electronics industry where the electrons and holes in a semiconductor display similar properties to a gaseous plasma.

There is thus great scope for applications of plasma physics and a resultant need for a more stringent definition of the criteria prescribing a plasma.

1.1. Debye length

The behaviour of a plasma is very different from a neutral gas because of the interactions between its charged particles. We will use the effects of these charges to define the criteria for a charged gas to be a plasma. First we will examine the concept of *quasineutrality*. Consider what happens when a ball of charge is introduced into a neutral plasma of ions and electrons. The plasma particles with opposite charge will move towards the ball until the potential difference vanishes and the plasma is overall neutral with a sheath of charge around the ball. To quantify this, we consider the effect of a potential ϕ_0 lying along the line $x = 0$ and calculate the general form of the potential, $\phi(x)$. We take the ions to be singly charged and infinitely more massive than the electrons and therefore fixed. The ion number density must be equal to the electron number density at infinity, $n_i = n_0$. Poisson's equation in one dimension is:

$$\varepsilon_0 \frac{d^2\phi}{dx^2} = -e(n_i - n_e). \tag{1.1}$$

We now substitute Boltzmann's relation, $n_e = n_0 \exp(e\phi/k_B T_e)$, where e is the electronic charge, T_e is the electron temperature and $k_B = 1.38\times10^{-23}\mathrm{JK}^{-1}$ is Boltzmann's constant, into equation (1.1). Applying a Taylor expansion, valid for the region where $|e\phi/k_B T_e| \ll 1$, i.e. away from the immediate vicinity of the line of charge, yields:

$$\varepsilon_0 \frac{d^2\phi}{dx^2} = en_0\left(\frac{e\phi}{k_B T_e} + ½\left(\frac{e\phi}{k_B T_e}\right)^2 + \cdots\right). \tag{1.2}$$

Retaining only the linear term on the right hand side, we see that the solution of equation (1.2) is given by:

$$\phi(x) = \phi_0 e^{-|x|/\lambda_D},$$

where

$$\lambda_D = \left(\frac{\varepsilon_0 k_B T_e}{n_0 e^2}\right)^{½}$$

is the *Debye length* which measures the thickness of the sheath of charge - the depth of plasma required to shield an externally applied potential - or the distance over which charge neutrality is not necessarily valid. Thus we may describe a plasma as quasineutral if the overall dimension of the plasma, L, is much greater than the shielding distance:

$$\lambda_D \ll L.$$

As a result of quasineutrality, the total ion and electron densities must be approximately equal so that $n_i \approx n_e \approx n$, where n is called simply the *plasma density*, but not so equal that local perturbations to this neutrality cannot take place.

1.2. Collective behaviour

Because a plasma is subject to electromagnetic forces, its particles interact with each other over great distances, not just when they undergo short-range interactions. Collisions are the only way neutral gas particles affect each other but every particle of a plasma is affected by every other one through the Coulomb force and so exhibits *collective behaviour*. One result of this collective behaviour is the shielding phenomenon. Debye shielding would be impossible if there were not enough particles to surround the introduced charge and so we must also require that the number of particles in the sphere of radius λ_D, $N_D = 4/3\pi\lambda_D^3 n_0$, is much greater than 1:

$$N_D \gg 1.$$

This limit may be considered to define a second criterion for a plasma.

1.3. Collision frequency

As we discussed above, plasmas interact through electromagnetic forces as well as collisions. In order to ensure that the gas behaves like a plasma and not simply as a neutral gas with collisions dominating, the frequency of typical collective plasma oscillations, ω_p, must be much higher than the frequency of hydrodynamic collisions with neutral atoms $1/\tau$. We therefore require:

$$\omega_p \, \tau > 1$$

which allows the electrons in the plasma to behave independently without being forced by collisions into complete equilibrium with the neutrals.
This condition is not independent of that in §1.2.

2. Two basic plasma frequencies

2.1. Plasma frequency

Plasma oscillations were originally postulated by Penning (1926) and were verified experimentally soon afterwards by Tonks and Langmuir (1929). We will consider what happens when a plasma is slightly perturbed from its equilibrium. In Chapter 3 we will derive the equations of continuity and momentum but for now we simply write down their linearised forms which are:

$$\frac{\partial n_e}{\partial t} + n_0 \nabla \cdot \mathbf{v}_e = 0, \tag{1.3}$$

$$\frac{\partial \mathbf{v}_e}{\partial t} = -\frac{e}{m_e}\mathbf{E}, \tag{1.4}$$

where we have again assumed the ions to be singly charged and stationary. $\mathbf{v}_e$ represents the electron velocity while m_e is its mass and we have taken the electron density to be the sum of the equilibrium density, n_0 and a small perturbed part, n_e. The motion of the charged particles leads to a perturbed electric field which can be related to the charge density using one of Maxwell's equations. We therefore have:

$$\nabla \cdot \mathbf{E} = -\frac{e n_e}{\varepsilon_0}.$$

Taking the time derivative of equation (1.3) and the divergence of equation (1.4) yields, on substituting for the electric field from above:

$$\frac{\partial^2 n_e}{\partial t^2} + \omega_{pe}^2 n_e = 0,$$

where the electron *plasma frequency* is given by:

$$\omega_{pe} = \left(\frac{n_0 e^2}{\varepsilon_0 m_e}\right)^{1/2}.$$

Thus $n_e \sim \exp(-i\omega_{pe} t)$ and the motion resulting from the small perturbation is a harmonic time variation at the plasma frequency. The plasma frequency is the natural frequency of vibration of a plasma.

Having defined the plasma frequency, we will now consider the passage of a plane harmonic wave through the plasma with all the perturbed quantities varying as $\exp i(\mathbf{k}\cdot\mathbf{r} - \omega t)$. In this way, we will demonstrate the significance of the plasma frequency to electromagnetic waves propagating in a plasma. From equations (1.3) and (1.4):

$$n_e = -\frac{i n_0}{\omega} \nabla \cdot \mathbf{v}_e \quad , \quad \mathbf{v}_e = -\frac{ie}{\omega m_e} \mathbf{E}.$$

The current flowing in the plasma due to the motion of the electrons is:

$$\mathbf{J} = -e n_e \mathbf{v}_e.$$

Finally, in order to construct a wave equation, we require Maxwell's equations:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad (1.5)$$

$$\nabla \times \mathbf{B} = \mu_0 (\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}). \qquad (1.6)$$

Taking the curl of equation (1.5), substituting from equation (1.6) and using the definition of the plasma frequency yields:

$$\nabla\times\nabla\times\mathbf{E} = i\,\omega\mu_0(-en_e\mathbf{v}_e - i\,\omega\varepsilon_0\mathbf{E})$$

$$= i\,\omega\mu_0(i\frac{n_e e^2}{\omega m_e} - i\,\omega\varepsilon_0)\,\mathbf{E},$$

$$\text{i.e.} \qquad k^2\mathbf{E} - (\mathbf{k}\cdot\mathbf{E})\mathbf{k} = \frac{\omega^2}{c^2}(1 - \frac{\omega_{pe}^2}{\omega^2})\,\mathbf{E}. \qquad (1.7)$$

For *transverse* oscillations (**k** perpendicular to **E**) the wave equation, (1.7), tells us that the *dispersion relation* for an electromagnetic wave propagating in an unmagnetised plasma has been modified from its free space form and is given by:

$$\omega^2 = k^2c^2 + \omega_{pe}^2.$$

For very high frequencies, $\omega \gg \omega_{pe}$, electromagnetic waves will propagate as if in a vacuum but at lower frequencies their behaviour will be altered until they reach a cutoff at the plasma frequency below which they cannot propagate. (The terms *cut-off* and *resonance* are used to describe points where the refractive index goes to zero and infinity respectively. Since the refractive index is directly proportional to wavenumber and inversely proportional to wavelength, cutoff corresponds to infinite wavelength and resonance to zero wavelength.) Applied fields of lower frequencies cannot penetrate the plasma because the more rapid plasma oscillations neutralise the applied field. Plasmas are therefore opaque to radiation of $\omega < \omega_{pe}$. In fact, this is not quite true. If the disturbing wave has a sufficiently high electric field, the electrons cannot shield out the effect of the wave completely and the wave can in fact propagate into the plasma. This phenomenon, related to the collisionless skin depth, c/ω_{pe}, thus allows a very large electric field to build up in the plasma, accelerating the electrons along it and raising them to relativistic energies. This acceleration process may be involved in producing cosmic rays and the relativistic electrons of radio sources in astrophysical plasmas (Tayler, 1982).

If **k** is parallel to **E**, (*longitudinal* wave) equation (1.7) reduces to:

$$\omega^2 = \omega_{pe}^2,$$

which represents the electron plasma oscillations discussed above. Because $\mathbf{k}\times\mathbf{E} = 0$, we may see from equation (1.5) that the perturbed magnetic field must also vanish - the wave is therefore also *electrostatic*. These plasma oscillations, often called Langmuir waves, do not depend on the wavenumber of the wave and so the group velocity of the wave, $d\omega/dk = 0$. The wave is therefore *stationary* and does not propagate in the limit which we have considered. In a real plasma, however, Langmuir waves will propagate to some extent due to two effects which we have not included. First, the oscillating electric field in a finite plasma extends beyond the region of the initial disturbance and so couples the oscillation into neighbouring layers of the plasma. Second, if the plasma has a non-zero temperature (and pressure), the velocity equation, (1.4) is modified and the dispersion relation for plasma waves gains a wavelength-dependent term:

$$\omega^2 = \omega_{pe}^2 + k^2 V_e^2 .$$

Here $V_e^2 = \gamma k_B T_e / m_e$ is the electron sound speed derived from γ, the ratio of the specific heats at constant pressure to constant volume, Boltzmann's constant and the electron temperature and mass. (The temperature characterises the width of the Maxwellian velocity distribution function and so is a measure of the mean kinetic energy of the random particle motion. In three dimensions, $E_{av} = 3/2 k_B T$ is the average kinetic energy possessed by the plasma particles.)

Langmuir waves are used as a diagnostic device by astrophysicists. In the early 1940's solar radiation with wavelengths of a few metres was observed causing interference with radio communications. Since then, this radiation has been studied in detail, classified into 5 main categories and its origins in the solar corona postulated. The so-called Type II radio bursts accompany solar flares and last some minutes, with the main spectral features drifting from low to high frequencies throughout their duration. Shklovsky (1946) was one of the first to suggest that these radio emissions were due to plasma oscillations. In this model, some source of excitation (possibly a shock wave) travelling out through the corona excites large amplitude plasma waves of successively lower frequencies, leading to the frequency drift in the observed spectra.

2.2. Cyclotron frequency

The motion of charged particles in a plasma is considerably altered by application of an external electric or magnetic field. The simplest context in which to examine this influence is orbit theory where the electromagnetic fields are taken to be known quantities and the fields created by the moving particles themselves are neglected.

A uniform, stationary electric field will simply serve to accelerate the particles along it. However a uniform magnetostatic field, $\mathbf{B}$ has a more subtle effect. Suppose that $\mathbf{B} = B\hat{\mathbf{z}}$. The equation of motion (which will be derived in Chapter 3) for a particle of charge q and mass m, moving with velocity $\mathbf{v}$ is:

$$m\frac{d\mathbf{v}}{dt} = q(\mathbf{v}\times\mathbf{B}). \tag{1.8}$$

Taking the components of equation (1.8) and writing $\dot{} \stackrel{\partial}{=} d/dt$:

$$m\dot{v}_x = qBv_y \qquad m\dot{v}_y = -qBv_x \qquad m\dot{v}_z = 0.$$

Differentiating the first two of these again with respect to time gives:

$$\ddot{v}_x = \frac{qB\dot{v}_y}{m} = -(\frac{qB}{m})^2 v_x, \tag{1.9a}$$

$$\ddot{v}_y = -\frac{qB\dot{v}_x}{m} = -(\frac{qB}{m})^2 v_y. \tag{1.9b}$$

Thus the parallel velocity is zero and both perpendicular velocity components vary harmonically with the *cyclotron frequency* defined to be always non-negative:

$$\Omega = \frac{|q|B}{m}.$$

We recognise that the solutions of equation (1.9) are:

$$v_{x,y} = v_\perp \exp(\pm i\Omega t + i\delta_{x,y})$$

where $\pm$ denotes the sign of the charge q, $v_\perp$ is the constant speed of the particle in the x,y-plane and the phase factors δ may be chosen in such a way that:

$$v_x = v_\perp e^{i\Omega t} \quad , \quad v_y = \pm i v_x.$$

Integrating these expressions and taking the real parts of the displacements, we find that the resulting motion of the particle obeys:

$$x = x_0 + r_L \sin\Omega t \quad , \quad y = y_0 \pm r_L \cos\Omega t, \tag{1.10}$$

where $r_L = v_\perp/\Omega$ is the Larmor radius of the orbit which has *guiding centre* (x_0, y_0). Equation (1.10) indicates that particles of opposite charge will gyrate in opposite senses. Also, since ions are much more massive and the Larmor radius is proportional to mass, the ion orbits will be much larger than those of electrons. The direction of gyration is always such that the magnetic field generated by the moving charged particles tends to reduce the ambient magnetic field - plasmas are *diamagnetic*.

When both electric and magnetic fields are present, the rotation round the magnetic field lines is modified by the electric field. If the electric and magnetic fields are not parallel, the particles drift under the influence of the non-zero **E**×**B** force acting perpendicular to the plane of both fields with a velocity of magnitude E/B. This drift velocity is independent of the mass and charge of the particle and so both ions and electrons will drift in the same direction. The three-dimensional orbit is therefore a slanted helix of changing pitch.

As the particles gyrate around the field lines, they emit radiation which may be used to infer the characteristics of the plasma. The line intensity is often used to provide information about the electron temperature whereas the width of the cyclotron emission spectrum may give evidence about the electron-ion collision frequency.

3. Applications of plasma physics

3.1. Astrophysical plasmas

In the universe, the plasma state is encountered almost everywhere but the variation in the plasma parameters - density and temperature - is immense as demonstrated in Figure (1.1). On the largest scale, over intergalactic distances of the order of 10^5 light years, space may consist of a low density plasma with a density of 10^{-1} to $10 m^{-3}$ while between clusters of galaxies there seems to be a gas of density $\sim 10^2 m^{-3}$. The interstellar medium inside a typical galaxy contains ionised hydrogen with $n \approx 10^6 m^{-3}$ although dense gas clouds may reach $n = 10^{12} m^{-3}$. Stars such as the Sun contain a wide variety of densities from $10^{15} m^{-3}$ in the solar corona and $10^{22} m^{-3}$ on the visible surface to $10^{30} m^{-3}$ in the stellar interior. Even higher densities are possible in such compact objects as white dwarfs and neutron stars. Of course, these changes in density do not occur at rigid boundaries and, in fact, the universe is inhomogeneous over all scalelengths.

Temperatures also vary dramatically - from 10^3K on the surface of the coolest stars to 10^{11}K in the central regions of ageing stars. Away from conditions of thermodynamic equilibrium, there is no uniquely defined temperature. When the particles of the interstellar gas have a Maxwellian distribution, the temperature is found to lie between 10^2 and 10^4K. The intracluster gas is observed from X-rays to have a temperature of the order of 10^8K. Over and above this, there are highly energetic constituents such as cosmic rays and emissions from various radio sources.

Finally, magnetic fields appear to occur throughout the universe - also with a wide range of variation. The general interstellar field is around 10^{-10}T with higher fields of 10^{-7}T existing in radio galaxies and pulsars having fields of 10^6 to 10^8T. Nearby, the Sun's average surface field is 10^{-4}T but far stronger fields (10^{-1}T) occur near sunspots. The Sun's field seems to be divided into regions of strong field and other regions of practically no field, as will be discussed in Chapter 4, and so the magnetic field of the Sun is highly inhomogeneous.

From this brief description, it is apparent that the properties of naturally- occurring plasmas are far from constant. As the plasma parameters change throughout the universe, all plasmas are to different degrees non-uniform and therefore far from the idealised description of the homogeneous theoretical models with which they are usually described.

3.1.1. Pulsars

Pulsars are distant objects, discovered by Hewish *et al.* (1968), which are of great interest to plasma physicists. They are observed through their emissions of pulsed high frequency radiation and it is from this radiation that their properties have been deduced. The shortest known period of a pulsar is 30ms for the pulsar in the Crab Nebula (although the recently-discovered optical pulsar (Kristian et al., 1989) in the supernova remnant, SN1987A, is thought to have a frequency of 1968 s^{-1}). If the period is the result of rotation, the pulsar must have an enormously high density to remain intact - a density comparable with a neutron star, around $10^{42}m^{-3}$. It therefore seems likely that pulsars are rotating neutron stars. Like most of the plasmas in the universe a pulsar's behaviour is governed by its magnetic field. Pulsar periods are known to increase gradually with time so that the pulsar rotation must be slowing down. This may be explained if the axis of its magnetic dipole moment is not aligned along the axis of rotation, which would result in a variable magnetic dipole radiating low frequency waves which would carry away angular momentum (Ostriker and Gunn, 1969). The observed variation in period can be explained if the pulsar magnetic field strength is of the order of 10^8T. It is thought that pulsars are formed in supernovae explosions when most of the star is ejected into space but a small fraction implodes leaving a high density neutron star or a black hole. If this view is correct and magnetic flux is conserved in the implosion, the magnetic field of an ordinary star could easily increase by a factor of 10^{10} which would be commensurate with the calculated field strength. The low frequency waves emitted would have sufficiently strong fields to pass through plasma and accelerate the particles to relativistic energies, as described before, and it is believed that it is the radiation from these relativistic electrons which we observe at radio frequencies. The radiation from pulsars may be used as a diagnostic tool in the study of the interstellar medium. Because of the large distance through which the waves travel, any small effect is enhanced and, for example, the dispersion of these waves may be used to estimate the density of the interstellar plasma. Similarly, the rotation of a polarised wave as it passes through an electromagnetic field (*Faraday rotation*) may be used to deduce the mean longitudinal component of the magnetic induction between the source and the observer.

3.1.2. Earth's radiation belts

Much closer to home, plasmas also play an important role in the region immediately beyond the Earth's atmosphere where charged particles from the Sun interact with the Earth's magnetic field. The thermonuclear reactions in the Sun's interior produce very high temperatures and maintain the entire Sun in a gaseous state - the divisions of the Sun's atmosphere will be discussed fully in Chapter 4 but our only concern here is how the changes in the solar atmosphere affect the Earth. The solar surface is an extremely turbulent region, presenting a constantly changing appearance to observers on Earth. Most of the dramatic activity occurring on the Sun is correlated to magnetic activity. *Sunspots* are dark regions of the photosphere with fields of $\sim 10^{-1}$T which are significantly stronger than their surroundings. They are not constant and grow, often coalescing with nearby sunspots into one large area which may persist for up to several months (Durrant, 1988). The number of sunspots also varies and possesses a regular cycle with a period of 22 years. These cycles have been associated with long-term weather patterns on Earth. Other solar activity which has a direct bearing on the Earth includes huge explosions on the solar surface called *solar flares* which are plasma jets accelerated up to 500km s^{-1}. Flares occur near sunspots and their frequency has also been related to that of sunspots. Both cosmic rays and visible radiation are increased by solar flares while particularly large flares may cause shock waves to travel through interplanetary space, compressing the Earth's magnetic field and leading to geomagnetic storms a few days later.

As a result of the supersonic expansion of the corona (the Sun's hottest, 2×10^6K, outermost layer) a stream of highly energetic particles, mostly composed of protons and electrons which are collectively known as the *solar wind*, is ejected into interplanetary space. The electron number density of the solar wind is around $5\times10^6 m^{-3}$ and particle velocities range between 300 and 600km s^{-1} - the exact values of both being closely linked to flare and sunspot activity. The magnetic field of the solar wind is $\sim 5\times10^{-9}$T. When the solar wind impinges on the Earth's dipole magnetic field, a collisionless shock is formed, causing a boundary known as the *magnetopause* beyond which most of the particles in the solar wind do not pass. Instead, they flow at supersonic speeds, round the outside of the magnetopause which is roughly spherical on the sunward side and cylindrical in the direction away from the Sun, as illustrated in Figure (1.2). As a result of compression by the solar wind, the Earth's magnetic field is also asymmetrical - with the field lines closer together on the side nearer the Sun and extended on the "downwind" side.

In the *magnetosphere* between the Earth and the magnetopause, the Earth's magnetic field lines form "magnetic mirror traps", confining charged particles. The magnetic moment,

$$\mu = \frac{\frac{1}{2} m v_{\perp}^2}{B},$$

of a particle is an *adiabatic invariant*; it remains constant for slow changes of the magnetic field in time or space. When a charged particle travels through an axially symmetric but slowly increasing magnetic field, μ must therefore be unaltered. Since B is increasing, the kinetic energy associated with its velocity perpendicular to the magnetic field must increase to compensate and this can only be done at the expense of the parallel kinetic energy. The parallel velocity will decrease and may vanish, in which case the particles will be reflected, hence the name magnetic mirror. If both ends of the magnetic field are configured in this way, the particle may be trapped indefinitely by repeated reflections and this is exactly what happens within the Earth's magnetic field lines. (The trapping is not perfect, however, and particles with sufficiently small perpendicular velocities lie in a *loss cone* in velocity space and are not confined.) The transit time for a particle to cross between two mirror points in the geomagnetic field is approximately 1s. As the particles follow the field lines, they also drift slowly azimuthally - electrons to the East and protons to the West - with their guiding centre tracing out a surface of rotation which will take about an hour for an electron of 40keV.

These regions of trapped particles, known as the *Van Allen* radiation belts, were discovered by early satellites whose radiation detectors failed at heights above 1000km but started functioning again when they returned nearer Earth (Massey, 1964). This was taken to be a sign of very high radiation levels in these regions. There are now known to be two main belts - one at a few thousand kilometres above the Earth and the second nearer 20,000km - composed mostly of electrons and protons but also including a small number of tritons and deuterons. The inner belt is attributed to cosmic rays penetrating the Earth's atmosphere and forming proton-electron pairs which are then trapped. Average proton energies in this band are 40MeV but are only of the order of 0.1 to a few MeV in the outer band. The outer band contains particles ejected from the Sun in periods of intense solar activity. Here, the electron density is approximately $n = 10^3 \mathrm{m}^{-3}$ and electrons in this outer belt typically have energies greater than 40keV.

At much lower heights, plasma effects are significant in the *ionosphere* which begins around 60km (below which there are no free electrons) and extends to over 10 Earth radii. The ionisation of particles in this region is again caused by solar radiation - in the extreme ultraviolet and X-rays. As the radiation penetrates deeper into the atmosphere it encounters a higher density of particles and a larger number of electrons is emitted, but the radiation is gradually absorbed and so less radiation will reach the lower layers. The outcome is a peak in electron production near

200km. The terrestrial magnetic field has a strong influence on ionospheric phenomena, such as radio communication (Budden, 1961). In the ionospheric polar regions, fantastic luminous displays - the *aurorae* - are caused by radiation from atmospheric species and particles of solar and cosmic ray origin being accelerated down the geomagnetic field lines.

3.2. Solid state plasmas

One rapidly expanding area of plasma physics research is that of solid state plasmas which are usually studied under the approximations of the *cold plasma* model, the model which we will use in Chapters 5 and 6. The degenerate electron gas in semiconductors, semi-metals and metals constitutes a plasma (Hoyaux, 1970). Although the number of particles in a Debye sphere of a solid is seldom large enough to make $N_D \gg 1$, quantum mechanical effects associated with the uncertainty principle can give some solids an effective electron temperature sufficiently high to satisfy this criterion for a plasma. Because of the lattice effects, the effective collision frequency is much less than would be expected for a solid with $n \approx 10^{29}\mathrm{m}^{-3}$, complying with the final plasma criterion described earlier. The study of solid-state plasmas began in the late 1950's and concentrated for a long time on wave propagation, particularly in semiconductors such as InSb. The crystal lattice of the solid can be described in terms of wave mechanics and the two species of the solid state plasma are electrons and holes, which are regarded as the positive charge carriers. In appropriate materials, free electrons and holes have been observed to exhibit similar oscillations and instabilities to gaseous plasmas. (For a comparison of the properties of solid state and gaseous plasmas, see Table (1.1)). Semiconductor holes may have effective masses as low as $0.01m_e$, leading to very high cyclotron frequencies even in moderate magnetic fields. Because of their relevance to semiconductors, solid state plasmas are most likely to prove useful in the field of electronic circuitry.

3.3. Thermonuclear fusion

Since the 1950's, much attention has focussed on plasma physics, not because of its numerous applications in astrophysics, but because of its potential in providing an almost limitless source of power through controlled thermonuclear fusion. When two light nuclei come close enough together, they fuse to form a new element with an accompanying release of binding energy in the form of kinetic energy of the reaction products. Of the many possible fusion reactions, the one of most interest in controlled nuclear fusion research is:

$$\mathrm{D} + \mathrm{T} \rightarrow {}^4\mathrm{He} + \mathrm{n} + 17.6\mathrm{MeV}.$$

where D represents deuterium, T is tritium, n is a neutron and He is helium. The main advantages of fusion power are

1) almost limitless, cheap and readily available fuels (seawater) ,

2) reduced environmental impact compared to other energy sources ,

3) less likelihood of misuse as a component of nuclear weapons.

In order for a fusion reaction to take place, the Coulomb repulsion between the two nuclei involved must be overcome and this is one reason why fusion reactions with the lightest elements are preferable. A beam of deuterons incident on a tritium target is unsuitable since most of the energy is then lost in ionising and heating the target and in scattering out of the beam. Also, colliding beams can never be made dense enough to produce more energy than that required to accelerate the particles in the beam. The solution is therefore to heat the particles to form a Maxwellian plasma in which the fast particles in the tail undergo fusion - *thermonuclear fusion.* A particle undergoing rapid thermal motion in a group of thermal particles has a greater chance of colliding and fusing than does a particle moving linearly along with a group of particles in a beam.

The minimum operating temperature for self-sustaining thermonuclear reactor is when the energy released by the fusing of nuclei just exceeds that lost from the plasma by radiation losses, mainly bremsstrahlung. This *break-even* temperature occurs for the D-T reaction at about 4keV , being an ideal limit. This temperature is calculated by assuming that no particles escape. However, neutrons are continually being lost from the plasma and their energy could be partly recirculated back to heat the plasma, but this would be costly.The alpha particles remain in the plasma, depositing their energy via collisions. *Ignition* occurs when the reaction rate is high enough to give the alpha particles alone sufficient energy to maintain the plasma temperature against radiation losses. Ensuring that the energy output is greater then the input imposes a condition on the plasma density, n, and confinement time, τ as well as on the temperature - a limit known as the *Lawson criterion* (Lawson, 1957). The minimum value of the product $n\tau$ required for the D-T reaction at $T \approx 13$ keV is $n\tau \approx 10^{20}\text{m}^{-3}$ s.

So far we have only discussed the positive side of controlled thermonuclear fusion as a source of power but there are, however, many difficulties to be overcome before a practical, large-scale fusion reactor can be built. These may be divided into three main categories:

1) Plasma Confinement ,

2) Plasma Heating ,

3) Fusion technology .

We will only be interested in the first two - the first because it helps to demonstrate realistic field configurations and the second because of its link with waves in inhomogeneous media.

3.3.1. Plasma confinement

Material walls could not withstand the high temperatures of the plasma and so an alternative method of confining the plasma is needed. Of even more importance is the fact that contact with the vessel walls would reduce the temperature of the nuclei to such a degree that fusion would be impossible. The majority of proposals for controlled fusion reactors have employed the charged nature of the ionised plasma to confine it using magnetic fields. The main approaches to confining the plasma for long enough to satisfy the Lawson criterion are:

1) Closed systems - toroidal systems which vary in the way they twist the magnetic field lines.

The internal plasma current in *tokamaks* produces a poloidal field in addition to the toroidal field. In *stellarators*, the twisting of the magnetic field lines is produced by external helical windings. *Multipoles* have their field lines primarily in the poloidal direction, produced by internal conductors and the toroidal field component of *spheromaks* is maintained entirely by plasma currents so that only poloidal field coils are required.

2) Open systems - magnetic mirrors

Mirror devices have axial magnetic fields to keep the plasma away from the walls and magnetic mirrors at both ends to trap the particles. Such machines therefore work on the same principle as the Van Allen radiation belts described earlier. One way of reducing particle losses from the ends is by connecting two mirror devices together via a central solenoid - the ‘tandem mirror’ concept.

3) Theta pinches

In these devices, a plasma current in the azimuthal direction and a longitudinal magnetic field produce a force which compresses the cross-sectional area of the plasma.

4) Laser fusion

Lasers are used to heat solid (cold) fuel element very quickly to very high temperatures. If the heating is fast enough, a large number of the nuclei will collide, fuse and release energy before the pellet has time to expand appreciably. Thus the nuclei are confined by their own inertia while they fuse and this is often described as *inertial confinement.* Such laser fusion studies comprise a branch of plasma physics in their own right and will not be considered further in this thesis.

Tokamaks and pinches

The plasma pinch phenomenon was first studied by Bennett (1934) who recognised that a plasma carrying a large current will contract radially. Suppose that a large current flows azimuthally round the exterior of a plasma cylinder, then this will cause a longitudinal magnetic field inside the plasma. This, in turn, induces an azimuthal current in the plasma in the opposite sense to that outside the cylinder. The resulting **J**×**B** force in the interior compresses or pinches the plasma until the magnetic pressure equals the kinetic pressure due to the particles' random thermal motion. A measure of the relative strengths of these two pressures is given by the plasma β:

$$\beta = \frac{p}{B_0^2/2\mu_0},$$

where p is the plasma pressure and B_0 is the value of the magnetic induction external to the plasma (equal to that of the plasma at the boundary). Plasma confinement schemes tend to work in either the very high or very low-β limit. Because the magnetic field strength is always less than B_0 and the sum of the magnetic and kinetic pressures is always constant, β is always less than 1 and the high β limit corresponds to $\beta \approx 1$. The device just described is the linear θ-pinch which is relatively stable whereas the z-pinch, where an axial current and an azimuthal magnetic field combine to produce the pinching effect, is subject to both the kink and sausage instabilities.

Linear pinches with the characteristics described above can be constructed with magnetic mirrors to reflect the particles but there will, of course, be losses at the ends and most of the particles will escape in a time $\tau \sim L/V_{si}$ where L is the length of the device and V_{si} is the ion sound speed. The simplest way of solving this problem appears to be to bend the pinch into a torus which is a closed system and will therefore not suffer from end losses. The main difference in closed systems is that currents cannot be provided by electrodes but must be inductively coupled.

Figure (1.3) shows a simple torus in which the lines of force are closed. The strength of such a toroidal magnetic field, which may be induced by winding current-carrying wires around the torus, decreases with increasing radius, $|B| \propto r^{-1}$. As a result, the ions and electrons have unequal Larmor radii in opposite halves of their orbits and so tend to drift to the top and bottom of the torus as shown, giving rise to an electric field. The interaction of the resulting electric field and the toroidal magnetic field produces an **E**×**B** drift outwards, of both ions and electrons, towards the walls of the container. Obviously, this is one of the first problems which must be overcome before equilibrium can be achieved in such a toroidal system. If there were a slight twist in the lines of force, so that the field lines were not closed over one circumference of the torus, particles would drift away from the centre of the plasma in one half of a circuit but towards it in the other half, thus "shorting out" the electric field. This particle drift may be counteracted by passing a current through the plasma. This plasma current leads to a poloidal magnetic field superposed on the toroidal field and results in helical field lines. The degree of twisting of the magnetic field lines is measured by the *rotational transform*, ι, which is the change in the angle of the magnetic field lines with respect to the minor axis in one turn around the major axis of the torus. (See Figure(1.4) for the geometry of a toroidal system.)

In addition to the two components of the field discussed above, toroidal systems also have a vertical magnetic field component to balance the natural tendency of a current ring to expand. Since the magnetic pressure of the poloidal field, $B_\theta^2/2\mu_0$, is larger on the inside, the major radius of the plasma tends to increase. The vertical field, $\mathbf{B}_v$, acts so that $\mathbf{J}\times\mathbf{B}_v$ is directed radially inwards. This field is often supplied by external coils but may be provided by transient eddy currents (image currents) induced in a copper stabilising shell surrounding the torus.

Figure (1.5) shows a suitable combination of externally applied fields for inducing the toroidal, poloidal and vertical magnetic fields in a tokamak. This is the basic design of large-scale tokamaks such as JET in England and JT-60 in Japan. Since the plasma current must be induced by a transformer, a tokamak cannot operate in a steady state and plasma is created and destroyed in each pulse of the transformer. To maintain stability in a tokamak, it has been found that the Kruskal-Shafranov limit, $\iota = 2\pi$, must not be exceeded. Helical displacement of the plasma as a whole is prevented if its wavelength exceeds the major circumference of the plasma, i.e.

$$2\pi a \frac{B_\phi}{B_\theta} > 2\pi R\,,$$

where a and R are the minor and major radii of the plasma and B_ϕ and B_θ are the toroidal and poloidal fields. This condition is equivalent to requiring $q = aB_\phi/RB_\theta > 1$, where q is the called the safety factor. (This is the limit for a circular cross-section.) In the presence of resistivity, the safety factor may have to be increased to avoid disruptive instability and generally $q > 3$.

The plasma β must be greater than 1% for economic reasons since the energy produced by a reactor is proportional to n^2 while the cost of the magnetic container increases as some power of the magnetic field. It is also desirable to make the plasma β as high as possible, to balance the kinetic pressure, since large magnetic fields are costly and difficult to maintain. Unfortunately, in a tokamak, the maximum value of β is limited to low values by the onset of magnetohydrodynamic (MHD) instabilities . Large tokamaks, such as JET, currently operate with β up to 10%.

MHD instabilities, which lead to wholesale motions where the plasma behaves like a conducting fluid, can also be eliminated by ensuring that the confining magnetic field passes through a minimum at some point in the plasma - the so- called minimum-B configuration. Tokamaks are in an average minimum-B configuration automatically since the poloidal field increases away from the minor axis and particles spend more time, on average, inside the magnetic well than outside. Microinstabilities, involving relative motion between the ions and electrons, may be reduced by altering the pitch of the helical field lines. Microinstabilities (e.g. drift waves) also occur despite this *shear* and may lead to anomalous transport.

The tokamak field is sheared because the toroidal field varies slowly across the cross section and the poloidal field has a minimum at the minor axis of the torus and a maximum near the outer edge. The magnetic field variation across a tokamak is illustrated in Figure (1.6a).

There are several difficulties with tokamak operation, mostly related to the equilibrium configuration. Plasma equilibria are far more complicated than for devices where all the confining fields are produced by external sources since changes in plasma properties lead to changes in toroidal current which then lead to changes in poloidal field which can alter confinement and the plasma parameters. Also, because the plasma is heated by Joule dissipation of the plasma current and this current is also needed to provide the rotational transform for equilibrium, the problems of confinement and heating cannot be treated separately. Tokamaks are relatively easy to construct because of their symmetry and because the complicated magnetic field geometry is produced by a simple arrangement of external coils. The plasma current required for the poloidal magnetic field has the additional effect of Ohmically heating the plasma - the problem of plasma heating will be discussed in the following section. They have been studied more extensively than most other confinement devices and, as a result, more is known about their optimum operating range than about any other. Perhaps this is why they, at present, seem to provide the best prospect for a full-scale working reactor.

Figure (1.6b) illustrates the distribution of toroidal and poloidal fields in a second type of toroidal pinch - the *reverse field pinch.* This device is characterised by a relatively weak stabilising toroidal field where $B_\phi \sim B_\theta$ so that the field lines spiral round the magnetic axis many times in going once around the torus. In addition, the toroidal field has the opposite sign at the centre of the plasma to its sign at the plasma's outer edge. Thus, the plasma must undergo a reversal of field near its axis and, at this point, the purely poloidal field is equivalent to the spatially rotating magnetic field which will be envisaged in Chapter 6. The result of this field reversal is a strong twisting of the field lines and this shear is primarily responsible for the stability of the reverse field pinch (RFP). Image currents induced in a toroidal shell encircling the plasma also help to stabilise the plasma when the shell is sufficiently closely fitting. This is a disadvantage in design which the RFP has over the tokamak. One particularly attractive feature of the RFP is that it can produce a very high ratio of plasma energy to magnetic energy, i.e. a high β, which is economically desirable, as argued above. In fact, the theoretical β limits for RFP's seem to be at least 3 times those for tokamaks.

Taylor (1974) showed that the reverse field configuration, at $\beta = 0$, corresponds to a minimum energy state in a flux conserving shell in the presence of a small amount of dissipation. This implies that the reverse field will be generated by the plasma and maintained even if it is not supplied from outside. This field reversal can be caused by motions similar to those which may produce the Earth's magnetic field. In practice, the reverse field can be established in a controlled fashion by fast programming the magnetic fields on microsecond timescales or on millisecond timescales by allowing the discharge itself to generate the reverse field. The latter process was employed on ZETA and other slow pinches even before Taylor explained the "self-reversal" in terms of plasma relaxation to a minimum energy state. After an initial unstable phase, the plasma in ZETA was observed to relax into a quiescent, stable state which might last up to $3ms$. The duration of this quiescent phase was limited by the decay of the reverse field due to the resistivity of the plasma. Proposals have been made to counter this decay using turbulence to regenerate the field but the turbulence may lead to unacceptably high losses. Alternatively, the pitch reversal might be achieved and sustained using currents flowing in external helical conductors similar to those used on stellarators.

The safety factor of the RFP need not be less than 1 since local MHD instabilities are suppressed by the strong shear. Thus the achievable values of B_θ and I_ϕ for a given B_ϕ are increased by a factor ~3 relative to a tokamak which implies that the required toroidal field is less in an RFP. If the losses in large RFP's were comparable to those in present day tokamaks, it might be possible to heat the plasma to ignition through Ohmic heating alone. There are also significant engineering advantages over tokamaks since the aspect ratio, a/R need not be minimised. A significant problem is that the basic mechanisms operating in an RFP are, thus far, less well understood than those in tokamaks.

Table (1.2) compares a set of toroidal pinch parameters which are of relevance to fusion for two tokamaks (JET and DITE) and two RFP's (ZETA and HBTX), the data being taken from the years indicated.

3.3.2. Plasma heating

The problem of raising the plasma temperature to that required for breakeven or ignition, and then maintaining that temperature over the lifetime of the plasma, has been tackled in several ways. The initial heating mechanism considered was Ohmic heating due to the current flowing in the resistive plasma. This type of heating occurs naturally in the two types of toroidal pinch described above but is unlikely to produce reactor temperatures because the plasma resistivity decreases rapidly with increasing temperature. Also, the plasma current in tokamaks is limited to maintain confinement and stability which restricts the power available from Ohmic dissipation. Of the many remaining suggestions for heating methods, two of the leading contenders are currently:

1) *neutral beam injection*, - the injection of high power beams of neutral hydrogen or deuterium atoms which can cross the containing field and are then ionised and trapped within the target plasma. This has proven to be the most successful method of plasma heating to date.

2) *radio frequency heating* - radio waves at a frequency matching a natural mode of the plasma excite a coupling resonance in the plasma.

We will restrict our attention to radio frequency (rf) heating because a great deal of work in the field of wave propagation in inhomogeneous media has been motivated by radio frequency heating in tokamaks.

The early analysis on rf heating near the ion cyclotron resonance extended the work of Alfvén (1942) and Åström (1950). As we will discuss in Chapter 5, a plasma can support a wide variety of wave motions and these provide a method of injecting external energy into the plasma, thereby heating it. There are four stages to rf heating - each one of which raises obstacles to overcome before the plasma can be successfully heated. First, there is the generation of the waves. Because the loss of power in the transmission between the generator and plasma is considered to be small, the equipment can be sited at a distance from the reactor, outside its high radiation area which allows access to the transmitters and is therefore one distinct advantage of rf heating.

Next, there is the question of wave penetration into the plasma. The resonance of interest may not be accessible to the incoming wave if the density of the plasma is too high, resulting in a cutoff where the wave will be reflected before the resonance is reached. It may, however be possible for the waves to *tunnel* through the region between the cutoff and resonance if this layer is not too wide (comparable to a wavelength) - a phenomenon akin to quantum mechanical tunnelling through a potential barrier. Alternatively, if the wave is incident from the inside of the torus, where the density is highest, it will meet with the resonance first. This situation is, of course, technologically much less convenient.

Third, the radio frequency waves must couple to a natural mode of the plasma and this process can involve several stages. When a wave is launched into a nonuniform plasma, its character will change as it propagates because the parameters of the plasma vary from point to point. In fact, if the parameters of the original wave are altered until they are sufficiently close to those of another possible oscillatory mode, some of the wave may convert into this second mode - the process of *linear mode conversion*. If the incident wave, or a wave into which it converts, encounters a resonance where the wavelength goes to zero, the transfer of electromagnetic energy to plasma wave energy can take place. This particular property of inhomogeneous plasmas - mode conversion - has motivated much of the work in rf heating in recent years (cf. the review article by Swanson, 1985). (There have also been parallel investigations of linear mode conversion in other fields of physics where wave propagation occurs in a non-constant medium - these will be discussed in detail in Chapter 2.) The wave in the inhomogeneous plasma is modelled using geometrical optics in which a WKB representation of the wave field is possible. Basically, this limits the investigations to "slowly-varying" media because the geometrical optics approximation is not appropriate near resonances where the plasma parameters vary greatly.

The final stage is to determine the absorption process for converting the plasma wave energy into thermal energy, a process which may take place in a variety of ways. The wave energy can be transferred directly to the plasma ions and electrons by collisions which randomise the directed particle velocity associated with the wave motion. Energy transfer may also be achieved through collisionless, or Landau damping, where certain particles resonate with the wave field and then lose their energy through collisions, or the particles may be heated via radiative dissipation.

As well as fundamental theoretical difficulties, rf heating is also technologically demanding because suitable antennae must be designed and built to produce the correct wave characteristics without overly affecting the plasma. Heating at several fundamental plasma resonance frequencies has been attempted including the ion and electron cyclotron frequencies (or multiples thereof), the lower hybrid frequency and fast Alfvén wave heating at twice the deuteron ion-cyclotron frequency. Ion cyclotron frequency heating was the first rf heating method attempted and has proved effective in a number of small plasmas. In this process, the waves propagate into a region of decreasing magnetic field, a *magnetic beach*, from below the ion cyclotron frequency where the wave energy is absorbed by the ions through ion cyclotron acceleration.

Experimental evidence indicates that the coupling of external waves into the plasma improves with increasing frequency so that the further development of sources to produce high power, high frequency waves may well be necessary in order to progress in this area. Questions also remain about the efficiency of rf heating to heat the core of the plasma and signs of impurities being heated in preference to the majority species.

4. Summary

In this chapter, we have given an outline of some of the basic attributes of plasmas in general and certain plasmas in particular which were chosen to provide relevant background information for the coming chapters. Having stressed the wide occurrence of plasmas in the universe, we proceeded to define a plasma to be a highly ionised gas satisfying the three main criteria:

1) $\lambda_D \ll L$,

2) $N_D \gg 1$,

3) $\omega\tau > 1$.

The derivations of the plasma and cyclotron frequencies, which will occur frequently in Chapters 5 and 6, were shown to arise naturally from the equations of motion for a plasma in two very simple sets of physical circumstances.

There followed sections on specific plasmas, highlighting the variety of plasma parameters which are observed, with the theme of inhomogeneity and variation constantly present. In particular, the magnetic field configurations described were selected to illustrate the wealth of structures possible and advantageous for either nature or man. The branches of plasma physics considered - astrophysics, solid state physics and fusion studies - are all being actively researched from different perspectives and are fruitful fields in their own right.

The investigation of wave propagation in inhomogeneous media, which is the theme of this thesis, was introduced in the context of supplementary heating for tokamaks but it could equally well have been introduced in relation to any one of several astrophysical problems e.g. solar coronal heating or the propagation of "starlight" through interstellar dust clouds. In the following chapter, we will examine the theoretical approach to mode conversion with particular emphasis on the shortcomings of many of the existing methods.

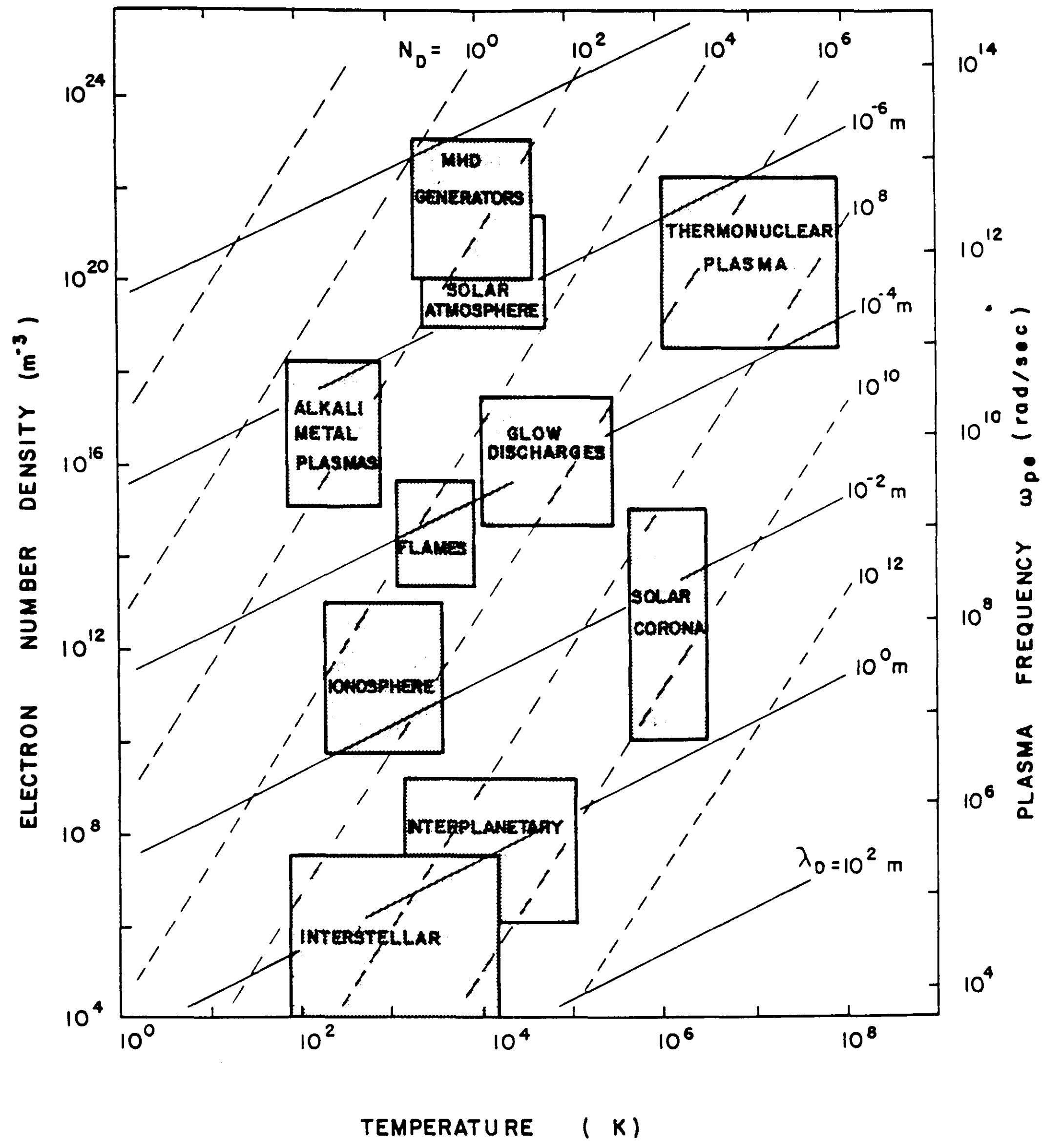


Figure (1.1) Ranges of temperature and electron number density for several laboratory and cosmic plasmas. Also indicated are the Debye length, λ_D, plasma frequency, ω_{pe}, and the number of electrons, N_D, in a Debye sphere.

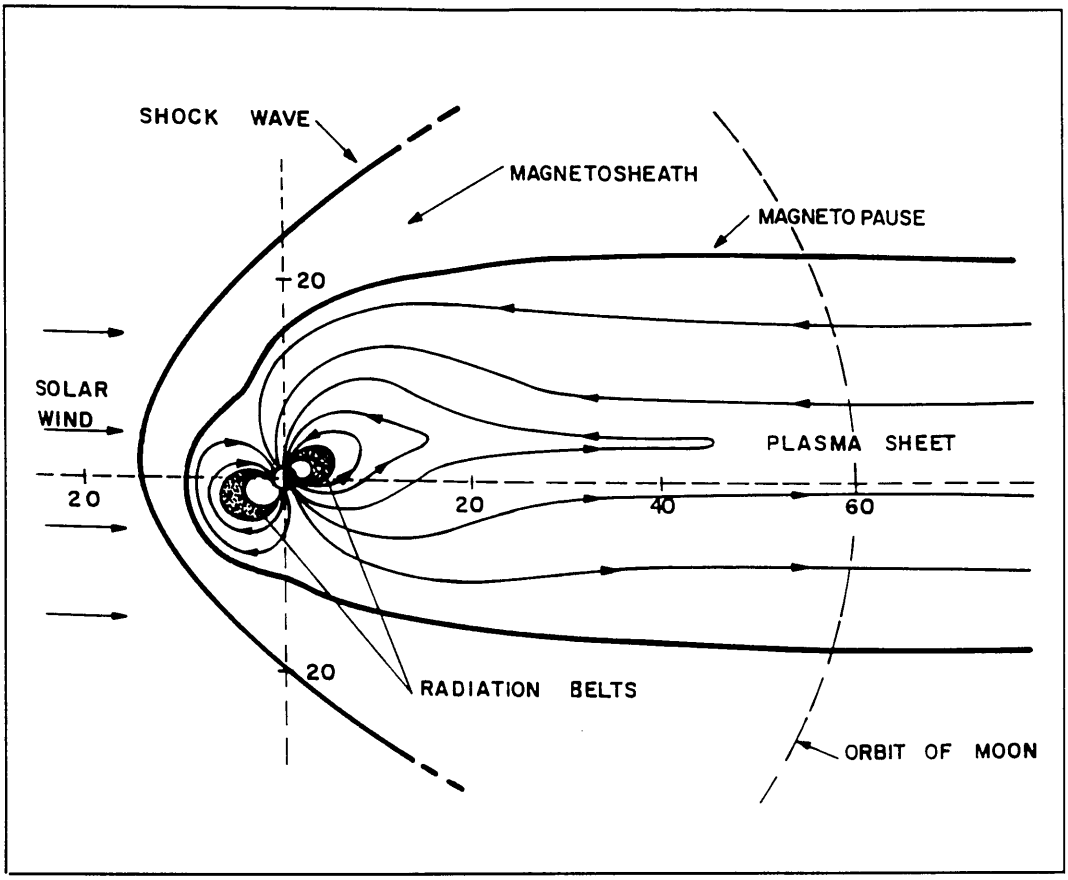


Figure (1.2) Schematic representation of the magnetosphere showing the *Van Allen* radiation belts and the turbulent region between the *bow shock* and the magnetopause - the *magnetosheath.* Distances are measured in units of Earth radii.

Parameter	**Unit**	**Gaseous Plasma**	**Solid-state plasma**
Plasma density	m^{-3}	10^{6} to 10^{24}	10^{18} to 10^{29}
Negative carrier mass	electron mass	1 (except negative ions)	10^{-3} to 1
Positive carrier mass	electron mass	10^{3} to 10^{5}	10^{-3} to 1
Plasma temperature	°K	10^{2} to 10^{9}	0 to 10^{3}
Plasma frequency	Hz	10^{3} to 10^{13}	10^{8} to 10^{15}
Debye screening length	m	10^{3} to 10^{-9}	10^{-5} to below lattice spacing
Negative carrier cyclotron frequency (for "usual" magnetic field)	Hz	0 to 10^{10}	10^{10} to 10^{13}
Positive carrier cyclotron frequency	Hz	0 to 10^{7}	same as for negative carrier
Dielectric constant at low frequencies		1	1 to 10^{3}
Average time interval between collisions	s	10^{7} to 10^{-10}	10^{-10} to 10^{-14}

Table (1.1) Comparison of parameters of gaseous and solid state plasmas.

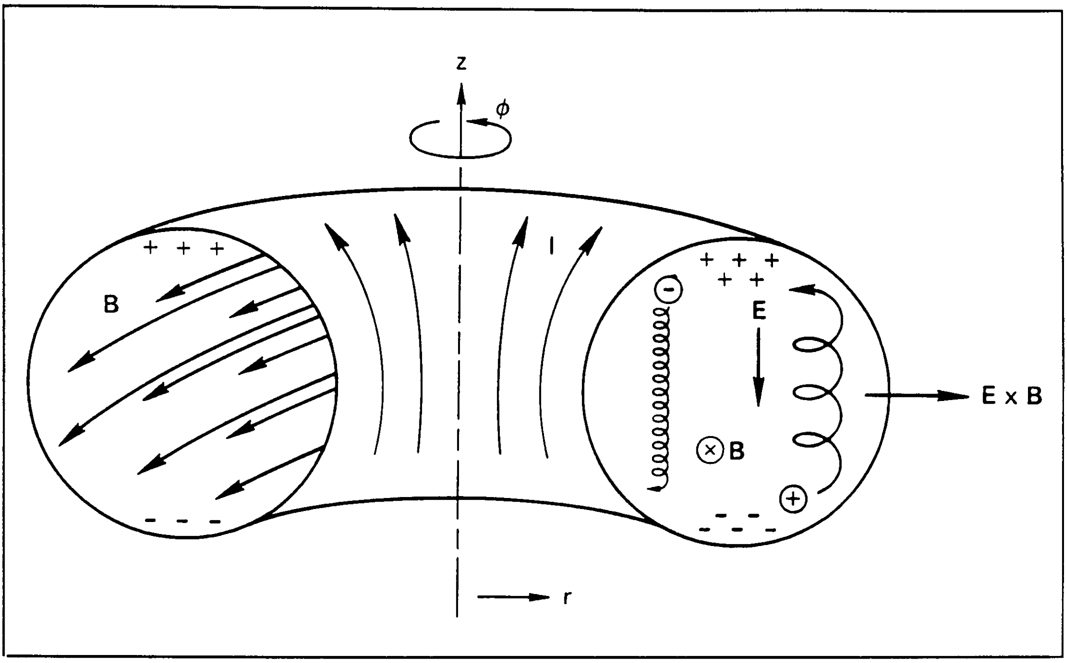


Figure (1.3) In a simple torus, in which the lines of force are closed circles, the magnetic field varies as $1/r$. The resulting ∇B drift causes a vertical charge separation, which in turn causes the plasma to drift outward.

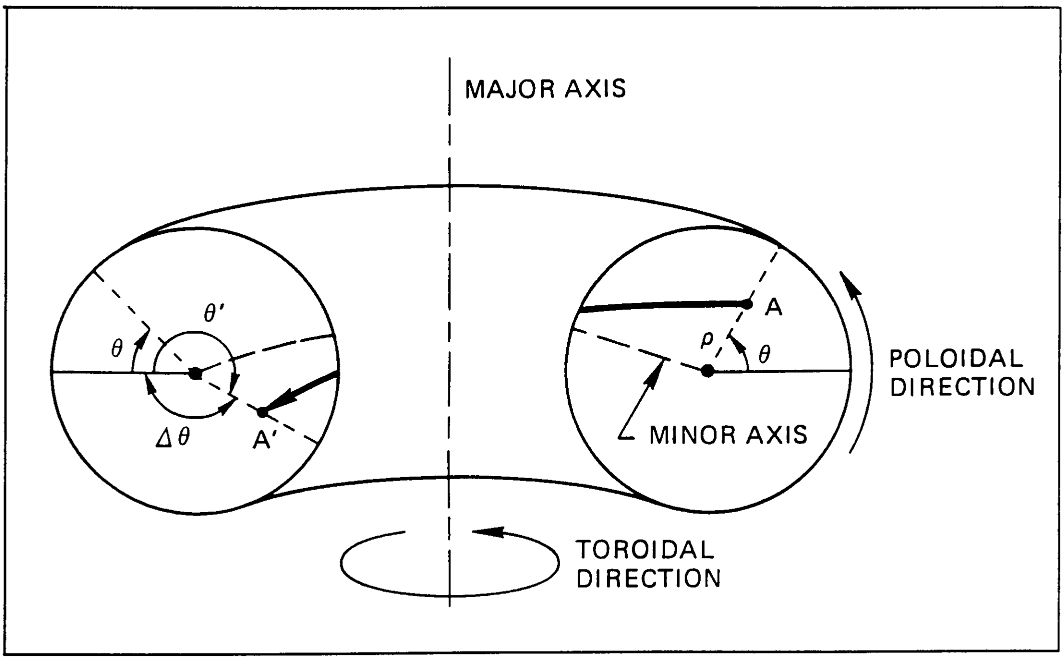


Figure (1.4) The geometry of a toroidal system with rotational transform. The field line from A–A' changes its azimuthal angle θ around the minor axis as it winds round the major axis.

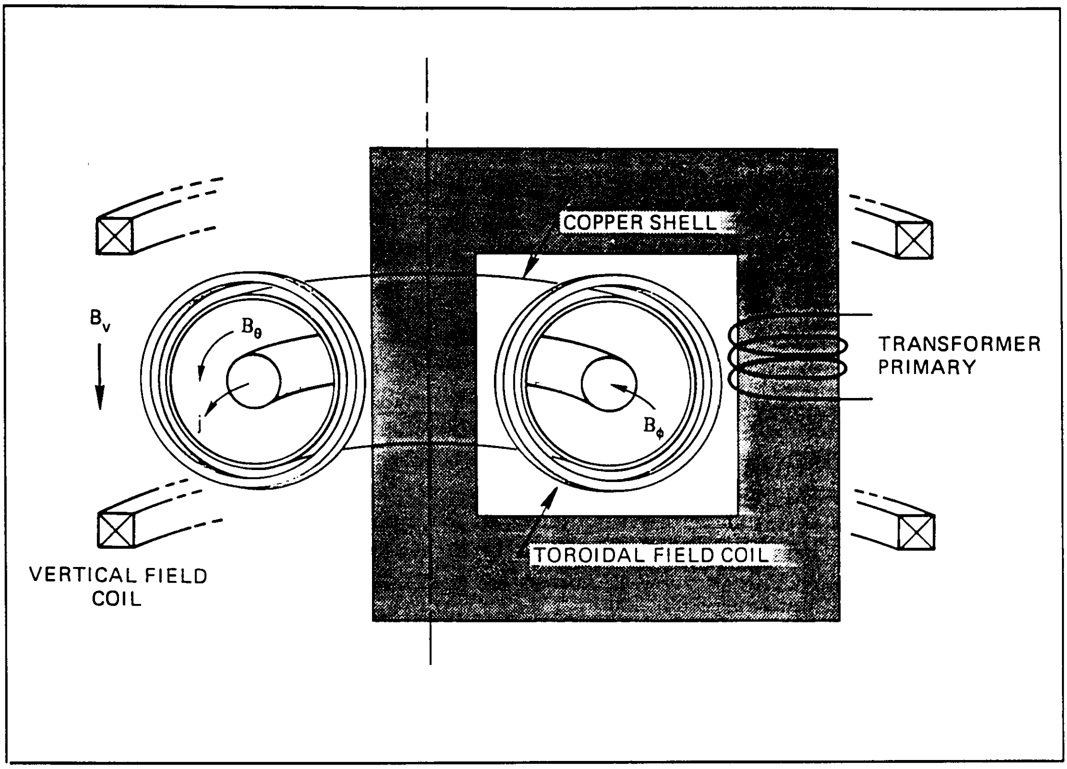


Figure (1.5) The toroidal field component, B_ϕ, is produced by current-carrying coils while the poloidal component, B_θ is produced by a large plasma current induced by a transformer. Additional stabilising forces are provided by a weak vertical field, B_v, and by eddy currents in a highly conducting copper shell.

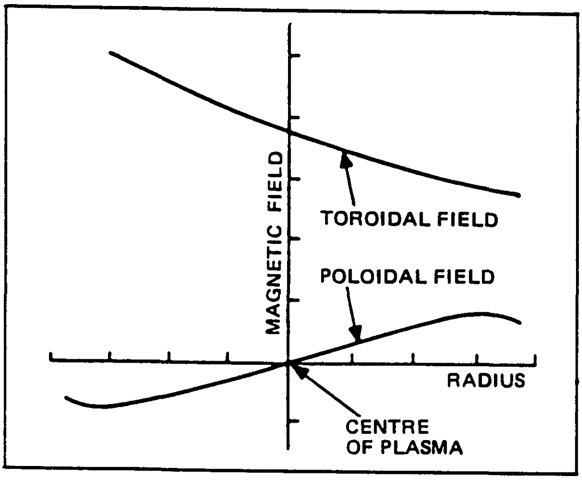


Figure (1.6a) Variation of the toroidal and poloidal magnetic field components across a tokamak.

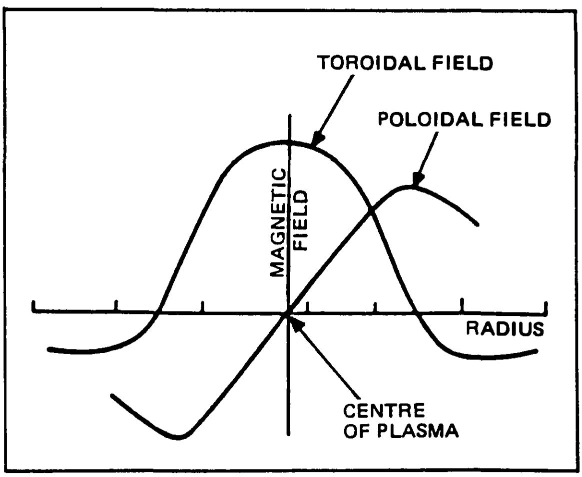


Figure (1.6b) Magnetic field variation across a reverse field pinch showing the field reversal in the outer region of the plasma.

Maximum Parameters *	**Tokamaks**		**Reverse Field Pinches**	
	JET(1989)	**DITE(1982)**	**ZETA(1969)**	**HBTX1(1980)**
Major radius (m)	2.96	1.17	1.5	1.0
Minor radius (m)	1.25, 2.10†	0.25	0.48	0.06
Current (MA)	6	0.28	0.5	0.3
Magnetic field (T)	3.4	2.7	0.08	0.3
Electron temperature (eV)	2.1×10^{4}	1.6×10^{3}	15	100
Electron density (m^{-3})	1.8×10^{20}	10^{20}	5×10^{19}	10^{21}
average β	0.11	0.02	0.1	0.3
energy confinement time (ms)	1200	40	10	0.4

* The maximum parameters quoted were not necessarily obtained all at the same time.

$\dagger$ For a non-circular cross-section, the two values given are for the horizontal and vertical minor radius respectively.

Table (1.2) Comparison of parameters of two tokamaks and two RFP's.

Chapter 2 - Mode Conversion

1. Motivation

When waves propagate in a homogeneous medium, the normal modes propagate independently of each other. However, the existence of an inhomogeneity in any of the plasma properties leads to a spatial variation of the coefficients describing the system. This, in turn, causes the modes of propagation to become position-dependent eigenstates, introducing the possibility of oscillations changing their character until modes which were originally totally dissimilar come to closely resemble one another. It is in this situation of a close matching of distinct wavemodes that we may expect *mode conversion*, i.e. the transfer of energy from one wave to the other.

Because most vibrating physical systems are to some extent inhomogeneous, mode conversion is of interest in many branches of physics. In particular, the majority of recent mode conversion studies have been motivated by the quest for additional heating methods in fusion research (Swanson, 1985). In a similar vein, solar physicists have attempted to explain the anomalous heating of the solar corona (Melrose, 1977) in terms of energy transfer between different types of wave while, nearer home, atmospheric physicists have found evidence for certain modes well above their theoretically predicted height and have used mode conversion to account for this phenomenon (Jones, 1970). These separate discussions share a common root in the intuitive notion that a redistribution of energy between the available oscillatory modes is most likely to occur wherever the properties of the modes involved are most similar and each theory has a graphical method (often a plot of frequency against wavenumber) for demonstrating the approach and divergence of modes in the particular region of interest.

It must be emphasised here that the term mode conversion refers to an interaction between linear waves - small perturbation expansions will be employed throughout this thesis and we shall not be concerned with non-linear phenomena, although much work has been concerned with such waves in recent years. As a consequence of the great interest in non-linear effects, the subject of linear wave coupling was comparatively neglected until the demands of the nuclear fusion community compelled theorists to consider the subject. Since then, several techniques have been suggested for heating the plasma to the critical ignition temperature above which fusion of the nuclei occurs. Originally, it was hoped that Ohmic dissipation of the plasma current would provide sufficient energy for heating large fusion devices like JET but it was soon recognised that additional sources of energy and alternative methods of depositing the energy in the plasma had to be sought. These questions have initiated many research topics in theoretical plasma physics and fusion- related technologies for almost three decades but the debate over the merits of the various techniques remains. A popular method for additional heating of Tokamaks is radio frequency heating, in which a radio signal directed into the Tokamak is converted to a mode which only propagates inside the plasma and which is ultimately damped, thereby releasing energy to heat the plasma ions. (Currently, the candidate wave for radio frequency heating is usually the compressional Alfvén wave.) The primary theoretical problems of radio frequency heating are concerned with the conditions under which the energy exchange occurs, where it occurs in the plasma and the partition of energy between the two waves involved - the fraction of energy transmitted, reflected and mode-converted.

Before examining any specific problems, we must discuss wave propagation in inhomogeneous media in general. One familiar way of studying oscillations within homogeneous media is via a dispersion relation. Because the properties of such a medium are constant, the wave's properties will also be constant and a linear wave of wavelength λ will always propagate with this wavelength. The relation between the wave frequency and wavelength, the dispersion relation, is an algebraic equation, obtained via Fourier transformation of the equations describing the system, and which may be used to identify uniquely the normal modes of the system. For a wave propagating in an inhomogeneous medium, such a relation cannot hold throughout the entire volume since the properties which give rise to the wavelength are not constant. Mathematically, this corresponds to the fact that the coefficients of the model equations are no longer constant and so the Fourier transform technique is no longer useful. Because the dispersion relation is not available in the inhomogeneous case, an alternative means of distinguishing the modes must be found and this new method of labelling modes must satisfy two basic conditions. First, it must incorporate the spatial variation present in an inhomogeneous plasma in a self-consistent and unambiguous manner and second, it must reduce to the dispersion relation in the homogeneous limit.

2. WKBJ theory

In this section, we consider briefly a powerful method for approximate solution of second order ode's with variable coefficients which was first suggested by Jeffreys (1923) and later modified by Wentzel (1926), Kramers (1926) and Brillouin (1926).

We note initially that the general first order linear differential equation,

$$y'+a(z)y = 0$$

(where $'$ denotes differentiation with respect to z) possesses an immediate solution in the form $y = \exp(-\int^{z} a(s)ds)$. On the other hand, a second order equation with variable coefficients such as:

$$a(z)y''+b(z)y'+c(z)y = 0 \tag{2.1}$$

has no such simple general solution. In general, the modes can no longer be identified in a similar manner - the parameter gradients play an important role and cannot be incorporated into such a straightforward approach. If, however, the coefficients of equation (2.1) vary "sufficiently slowly" a good approximation to the solution is given by an expression similar to the solution of the first order equation and this is the basic assumption of WKBJ theory. It is useful to have equation (2.1) expressed in normal form for what follows, and it may be rewritten:

$$y''+f(z)y = 0, \tag{2.2}$$

where $f(z)$ is often referred to as the wave potential, by analogy with quantum mechanics. For a constant potential, f_0, the solutions are $\phi = e^{\pm i\sqrt{f_0}z}$ so that the motion is oscillatory for a positive potential but evanescent for a negative one. The roots of the potential (the transition points) indicate where fundamental changes in the nature of the solutions take place and hence denote points where the oscillation will be radically altered - either reflected or absorbed. In this homogeneous case, the wavenumber of the oscillation is related to the potential, as shown above, by $k^2 = f_0$ and points where the potential vanishes are points where the wave is reflected whereas infinities in the potential indicate resonances where wave energy is absorbed.

If we attempt a solution to equation (2.2) of the form $y = \exp(\pm i \int^{z} k(s)ds)$, which is a simple extension of the solution for constant coefficients, there is an error term, $\pm i\, k'\exp(\pm i \int^{z} k(s)ds)$, left over. If the gradient may be taken to be small so that the potential is slowly varying, this is a reasonable approximation to the true solution and is, in essence, the WKBJ theory. The procedure may be further sophisticated by making the amplitude a function of position so that our new estimate of the solution is $y = A(z)\exp(\pm i \int^{z} k(s)ds)$. Substituting in equation (2.2), we find that on neglecting the second derivative, the wave amplitude must satisfy:

$$2kA' + k'A = 0 \quad \text{or} \quad A \propto \frac{1}{|k|^{1/2}}.$$

The second derivative may be neglected if it is legitimate to assume that f, and hence k, is a slowly varying function of position. It is obvious that near the transition points, the amplitude does not satisfy this criterion and these approximate solutions no longer apply. The WKBJ theory therefore provides formulae for the asymptotic solutions but says nothing about the behaviour of the function near the transition points. The solution is sure to fail at the zeroes of the potential where the approximation of a slowly varying medium no longer applies. In order to calculate the amount of each independent solution in one asymptotic region, when their ratio is known at the other extreme, the solutions must be connected through the reflection region. In order to do this, the solution is analytically continued into the complex plane and a connection formula calculated using the pertinent boundary conditions.

WKBJ theory states the results which we have just deduced in more mathematical language. It is designed to solve ode's of the type:

$$y'' + h^2 f(z,h) y = 0$$

where the parameter h is taken to be large and positive and for which the following conditions are satisfied (Heading, 1962):

1) $f(z)$ must be continuous for all z.

2) $f(z,h) \rightarrow const.$ as $h \rightarrow \infty$ for fixed complex z.

3) $|f'/f|^2 \ll 1$ and $|f''/f| \ll 1$ for all $|z| > z_c$ for some z_c.

The approximate independent solutions are then:

$$y_{1,2} = [f^{-1/4} \exp(\pm ih \int^{z} f^{1/2} ds)][1 + O(\frac{1}{h^2})]$$

where $O(\frac{1}{h^2}) \sim 5f'^2/16f^2 - f''/4f$.

3. "Local dispersion" theories

Over small regions of inhomogeneous plasma, if the plasma parameters vary sufficiently slowly, the coefficients of the system equations have approximately constant values, the terms involving derivatives of the parameters are small and a *local dispersion relation* is appropriate. A major advance in studying mode conversion came with Stix's (1965) work on radio frequency heating of tokamaks near the lower hybrid frequency where he used the local dispersion relation to reconstruct a differential equation with variable coefficients. The asymptotic solutions of the resulting equation could then be found using suitable integral-transform or phase-integral methods. In order to return to an ordinary differential equation from the dispersion relation, the inverse Fourier transform,

$$ik \rightarrow \frac{d}{dz}$$

must be taken (where z is the direction of inhomogeneity). Because it is based on extending a simple algebraic equation, this approach avoids the often formidable algebra involved in deriving a complete vector equation from the basic evolution equations for the system variables when the magnetic field, density etc. are functions of position. (The latter procedure will be applied, in Chapters 4 and 6, to two different physical models, illustrating the quantity of algebra required to construct the problem in this way. Although considerably slower to formulate, this method results in the only self-consistent statement of the problem.) Using the inverse transform of the local dispersion relation has therefore proved a very popular method, with all the predominant mode conversion theories being based on this principle, although it has been considerably developed.

The use of such inverse Fourier transform techniques has, however, several inherent weaknesses. First, the local dispersion relation is only applicable over regions where the coefficients are slowly varying and this does not hold near mode conversion points which are dependent for their existence on gradient effects and which are absent from uniform media. By re-introducing variable coefficients in the latter stages of analysis instead of including them self- consistently from the start, all information about gradients in the zero-order quantities is lost. The fact that the derivative terms cannot be incorporated naturally into such a scheme was recognised by Stix (1962). His more recent work (Stix and Swanson, 1983) has instead concentrated on developing analytic solutions for ordinary differential equations which can then serve as comparison equations for those arising in mode conversion studies. The technique aims to solve specific fourth order ode's by forcing the given ode into a general standard, soluble form.

A second difficulty with the inverse Fourier transform technique is highlighted by a major difference between the strategy of the two remaining theoretical camps. When the ode is reconstructed in the manner described above, how is the dependent variable to be assigned? Cairns and Lashmore-Davies (1983) consider the ode in the wave amplitude whereas Fuchs, Ko and Bers (1985) employ an equation describing the power flow. In the theory of Fuchs, Ko and Bers, the system may possess several allowed modes of oscillation (leading to a fourth or higher order dispersion relation) and these are again identified with the roots of the local dispersion relation and mode conversion is defined to be the redistribution of energy between these roots. From this basic premise, a complicated formalism is developed and an embedded dispersion relation describing the propagation of only two modes is extracted, in order to generate coefficients of transmission, reflection and mode conversion when this pair of waves interacts. The points where such interaction takes place are given by a set of conditions which must be satisfied simultaneously. Using $\hat{D}(k,z)$ to represent formally the embedded dispersion relation, the authors identify mode conversion points to be those critical points which are saddle points of the mapping $w \equiv \hat{D}(k,z)$, satisfying

$$\hat{D}(k_c,z_s) = 0 \quad , \quad \frac{\partial \hat{D}(k_c,z_s)}{\partial k} = 0 \quad \text{and} \quad \frac{\partial^2 \hat{D}(k_c,z_s)}{\partial k^2} \neq 0$$

which restricts the branch points of interest to pair-wise coupling events. The resulting ode to be solved is, in normal form:

$$\frac{d^2\phi}{dz^2} + Q(z)\phi = 0$$

with the potential, Q, given by $Q = -2(\partial \hat{D}(k_c,z_s)/\partial k) / (\partial^2 \hat{D}(k_c,z_s)/\partial k^2)$.

Although this theory appears to be mathematically sophisticated, dealing largely with branch cuts and saddle points in phase space, its physical basis is again unsound. The authors take the local dispersion relation as their starting point and so their arguments contain a flaw at the most fundamental level. It was only possible to construct the dispersion relation initially because the parameters contained no spatial variation and adding, at a later date, gradient terms which have been neglected in the first instance is not justified.

A further inconsistency which is apparent is that, although the local dispersion relation must be regarded at best as a local phenomenon, the authors employ asymptotic expansions to calculate the power flow. This gives the dispersion relation a global significance which it cannot merit and may result in a departure from the domain of the problem.

Cairns and Lashmore-Davies tackle the problem from a different viewpoint - that of a pair of coupled equations describing the characteristics of two modes of the plasma - but still rely on the inverse Fourier transform of a local dispersion relation as the basis of their treatment. They illustrate graphically the possibility of mode conversion in regions of (k_z, z) space where the roots of the dispersion relation come close together and then diverge. In these regions, they suppose that the dispersion relation may be approximated by:

$$(\omega - \omega_1)(\omega - \omega_2) = \eta$$

where $\omega_{1,2}$ are the frequencies of the uncoupled modes and η is the coupling function which is only significant in the neighbourhood of the coupling point. η consists of the remaining terms of the local dispersion relation when it has been factorised as shown. Taylor expanding this approximate dispersion equation about the assumed mode conversion point, yields a local dispersion relation:

$$(ak - ak_0 + b\xi)(fk - fk_0 + g\xi) = \eta_0$$

where a, b, f, g are partial derivatives arising from the expansion about the point (k_0, z_0), $\xi = z - z_0$ and η_0 is the value of η at the critical point. In their early work (Cairns and Lashmore-Davies, 1982), the authors performed the inverse Fourier transform at this stage but a straightforward replacement of k by the operator $-id/d\xi$ is ambiguous, as mentioned earlier, and the differential equation which results does not satisfy energy conservation. If however two wave amplitudes, $\phi_{1,2}$, are introduced, the mode conversion process may be viewed as a result of the coupling between these two waves and the local dispersion relation may be divided equally between the appropriate pair of ode's:

$$\frac{d\phi_1}{d\xi} - i(k_0 - \frac{b}{a}\xi)\phi_1 = i\lambda\phi_2,$$

$$\frac{d\phi_2}{d\xi} - i(k_0 - \frac{g}{f}\xi)\phi_2 = i\lambda\phi_1.$$

where $\lambda^2 = \eta_0/af$. Note that the coupling function has been divided equally between the two modes and that the extent of the coupling is not dictated by variation of the plasma parameters but by the magnitude of the remainder terms from the local dispersion relation. The conservation of energy flux is resolved by noting that the sum of the squares of the wave amplitudes remains constant for all z. In order to complete the solution of the problem, one of the wave amplitudes must be eliminated and the final equation may be manipulated into the form of the Weber equation for which the asymptotic solutions are known functions.

The fundamental basis is the same as for the other theories outlined above, since Cairns and Lashmore-Davies rely on being able to make the dispersion relation a function of position to generate their differential formulation of the problem. Although the concept of a pair of coupled modes is appealing, the coupling term itself exhibits certain undesirable characteristics. It does not result from physical effects but rather from the remaining terms of the dispersion relation when the desired approximate form has been factored out. Instead of being heavily dependent on position around the coupling point, it is constant here.

Recently, there has been growing criticism of the inverse Fourier transform method, and the resultant omission of parameter gradient terms, in the light of evidence from alternative solutions (Friedland, 1986) that the missing terms significantly alter the final mode conversion ratios. All in all, a more consistent description of wave propagation in inhomogeneous media is required. An alternative approach is to construct the vector differential equation describing the system directly from the fundamental model equations without attempting to generalise the dispersion relation which is only truly appropriate to a uniform system. In this way, all the information about parameter gradients in the background quantities is included in the equilibrium statement of the problem, which is an integral part of the linearisation process. In order to construct a scalar ode from this full equation, it may be necessary to apply certain approximations but, because these are approximations applied to an exact specification of the problem and not to an approximate one, this is still greatly preferable to using the inverse transform method. If the only necessary approximation occurs at the final stage of solving the equation, an estimate of the accuracy of the technique may be made, whereas the error involved in the many approximations made in the works discussed above cannot be evaluated.

4. Alternative approach

For mode conversion to be possible, there must be at least two distinct types of wave motion allowed in the medium. The only mode conversion possible in the simplest case of a single wave which is free to propagate in one of two antiparallel directions is simply reflection, whereas for more complicated systems which can support more types of wave, conversion between different categories of wave may also be possible. Because of the interest in mode conversion as a method of depositing the energy required for fusion, plasmas have provided the usual medium for mode conversion studies, as evidenced by the wealth of plasma physics literature devoted to this field over the last few years. Plasmas are ideal candidates for such studies but are by no means the only kind of medium in which waves may alter their character in this way. Indeed, any physical system which is free to oscillate under the influence of a number of applied forces will exhibit a larger number of eigenstates and might therefore sustain mode conversion. Waves in a compressible atmosphere under the influence of gravity are acted on by both the gravitational force and a restoring force due to the buoyancy force. They therefore fulfil this basic requirement for mode conversion and, because no magnetic forces are involved, the geometry of the system (and hence the algebra involved in its analysis) is greatly simplified from the plasma case. (These waves are discussed in detail in Chapter 4.)

Having introduced the problem from a physical standpoint, we must now find a unique and unambiguous mathematical way to label the modes which propagate in the inhomogeneous plasma. The properties of the modes must vary with position throughout the plasma in a manner dictated by the variation of the physical variables such as magnetic field strength and density. We follow the argument of Clemmow and Heading (1954), parts of which have been employed by authors in the past, some of whom do not seem to have appreciated it fully. In Heading's method, equation (2.1) is viewed as a pair of coupled equations and this aspect of its nature is fully exploited in order to arrive at a description of the modes and the degree of coupling between them. Equation (2.1) may be written as a system of first order equations by considering the column vector containing y and its first derivative:

$$\mathbf{y}' = \mathbf{M}\mathbf{y} \tag{2.3}$$

where

$$\mathbf{y} = \begin{bmatrix} y \\ y' \end{bmatrix} \quad \text{and} \quad \mathbf{M} = \begin{bmatrix} 0 & 1 \\ -c/a & -b/a \end{bmatrix}.$$

The eigenvalues of this matrix may be considered to be the eigenvalues of equation (2.1) and are given by:

$$\lambda_{1,2} = -\frac{b}{2a} \pm \left(\frac{b^2}{4a^2} - \frac{c}{a}\right)^{1/2} . \tag{2.4}$$

We shall use these eigenvalues to distinguish the natural modes of equation (2.1) in the same way that wavenumbers would represent the waves in homogeneous media. To demonstrate this equivalence, we can construct the pair of coupled first order ode's which correspond to equation (2.1) and will show that the coupling and propagation of the new dependent variables is inextricably linked to the eigenvalues defined above. In the limit of constant coefficients, we will show that the eigenvalue equation reduces to the form of the dispersion relation describing the homogeneous wavenumber.

We now perform the transformation of variable $\mathbf{y} = \mathbf{A}\mathbf{u}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} \quad \text{and} \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

$\mathbf{A}$ is the matrix of eigenvectors corresponding to the eigenvalues defined by equation (2.4) and so may be used to diagonalise $\mathbf{M}$. On substitution in equation (2.2), we obtain

$$\mathbf{u}' = \mathbf{A}^{-1}\mathbf{M}\mathbf{A}\mathbf{u} - \mathbf{A}^{-1}\mathbf{A}'\mathbf{u}$$

which may be expressed solely in terms of the eigenvalues as

$$\begin{bmatrix} u_1' \\ u_2' \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \frac{1}{(\lambda_1 - \lambda_2)} \begin{bmatrix} -\lambda_1' & -\lambda_2' \\ \lambda_1' & \lambda_2' \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

In the case of constant coefficients, the diagonal matrix of eigenvalues would remain but the second matrix containing the derivatives of the eigenvalues would disappear. The latter may thus be viewed as a coupling matrix describing the interrelation between the eigenstates caused by the variation of the physical parameters. From the elements of the coupling matrix we can see that the two most important factors affecting the coupling are the magnitude of the derivative of the eigenvalues and the separation of the eigenvalues. This agrees with our physical intuition that coupling is strongest where the variation is greatest and where the modes are most similar. Thus we have a mathematical formalism with which to express the physical behaviour already described. Separating the vector equation into its constituent parts yields a pair of coupled equations:

$$u_1' + \left(\frac{\lambda_1'}{(\lambda_1 - \lambda_2)} - \lambda_1\right) u_1 = -\frac{\lambda_2'}{(\lambda_1 - \lambda_2)} u_2 \tag{2.5}$$

$$u_2' - \left(\frac{\lambda_2'}{(\lambda_1 - \lambda_2)} + \lambda_2\right) u_2 = \frac{\lambda_1'}{(\lambda_1 - \lambda_2)} u_1.$$

It is interesting to observe that the coupled equations employed by Cairns and Lashmore-Davies are of the same form as equations (2.5) except that the coupling function in their case arises from remainder terms in the local dispersion relation and not from parameter gradients which are an integral part of the statement of the problem. The coupling coefficients of equation (2.5) possess a certain symmetry, but are not equal as Cairns and Lashmore-Davies' coupling factors are assigned to be. In the appendix, it is demonstrated that the above analysis may be extended to equations of third and fourth order and that in these cases the coupling between the modes is again in direct proportion to the gradient and inversely proportional to the separation of the eigenvalues which are interacting.

For negligible gradients in the eigenvalues, the right hand sides of equation (2.5) go to zero, the above pair of equations decouple and each has the trivial solution:

$$u_i(z) = \exp\left(\int^{z} \lambda_i(s)\, ds\right).$$

Thus λ is equivalent to a wavenumber (differing by a phase factor, i) in the limit of constant coefficients and we have shown that the eigenvalue description reproduces the dispersion relation in the uniform case. We will henceforth use this method to identify the eigenmodes of the inhomogeneous medium.

The eigenvalue description does not in itself provide a solution to the problem but is useful as a method of labelling the eigenvalues even in regions where the potential vanishes and the amplitude of the WKBJ approximation becomes singular. Because it reduces to the familiar dispersion relation in the homogeneous limit, it provides a mathematical description of the modes to accompany the purely intuitive one. One of its great strengths lies in the fact that it includes the coupling between different modes directly from the physical variables of the system and not as an external factor which is included ad hoc at a later stage of the analysis.

The full solution of the problem can only be obtained, however, by solving the original ode, equation (2.1), using the technique described in §2. The ode is reduced to a standard form, say Whittaker's equation or Weber's equation for which the asymptotic solutions are tabulated, but failing this, numerical solution may be necessary since analytic solutions are only known for a limited set of standard second order differential equations. This is one of the fundamental difficulties in mode conversion studies and much effort has been spent in trying to circumvent it. One of the strongest arguments in favour of using the methods discussed in §3 is that the differential equations obtained are in, or can be easily manipulated into, the form of a standard equation such as the Weber equation. When the system's evolution equations are combined to produce the final equation, there is no such guarantee that it will fit neatly into a class of analytically soluble equations. We will be confronted with precisely this problem in Chapter 4.

It must also be noted that equation (2.1) is in general derived from a set of partial differential equations and, in order to be able to reduce the problem to a single ode at all, we must assume that the parameters vary with one independent variable only. Finally, there is the problem of higher order ode's. Equations of the type of (2.1) may be solved analytically only if they fall into particular classes with coefficients obeying strictly prescribed rules. If a medium can support more than two wave modes, it will often be described by a fourth or higher order ode; many problems involving inhomogeneous plasmas fall into this category. Physically, this means that there are potentially more channels available for mode conversion. Mathematically, such equations are even more difficult to solve, with fewer standard methods known. Stix and Swanson have examined fourth order equations exclusively while Fuchs et al. have attempted to address the problem of reducing the order of the differential equations, but their final recipe for finding the mode conversion points and relevant coefficients is only applicable to the case of pair-wise coupling of modes.

5. Eigenvalues and potential

The eigenvalues defined in the previous section have a close link with the wave potential. The transformation which reduces equation (2.1) to normal form, requires the definition of a new dependent variable, $\phi = y\exp(\tfrac{1}{2}\int^{z} b/a\, ds)$, to obtain

$$\phi'' + (\frac{c}{a} - \frac{(ab'-a'b)}{2a^2} - \frac{b^2}{4a^2})\phi = 0.$$

The general form for the potential is then given by

$$f = -(\frac{b^2}{4a^2} - \frac{c}{a}) - (\frac{b}{2a})'. \tag{2.6}$$

As was discussed in §2, a constant potential, f_0, gives rise to either oscillatory or evanescent modes depending on the sign of f_0. When f_0 goes from positive to negative to positive, the wave must be oscillatory on either side of an evanescent layer. In quantum mechanics, the potential function is given by the difference between the kinetic energy of the wave packet and the background potential and such behaviour is therefore traditionally called a *potential barrier*. If however, the behaviour of f_0 is negative-positive-negative, this is analogous to a quantum mechanical *potential well* in which oscillatory solutions are only possible in the central region. (Often, the application of suitable boundary conditions to this case results in only a discrete set of possible eigenvalues and not an infinite spectrum of normal modes.) Quantum mechanical studies of wave scattering by potential wells and barriers are thus simply a subset of possible mode conversion events.

In the inhomogeneous case, the potential may be rewritten in terms of the eigenvalues, using the definition of the eigenvalues from equation (2.4) and the definition of f (equation (2.6)), as:

$$f = -\tfrac{1}{4}(\lambda_1 - \lambda_2)^2 + \tfrac{1}{2}(\lambda_1 + \lambda_2)'. \tag{2.7}$$

It is interesting to note that the important quantities are again the separation of the modes and the magnitude of the derivatives which were also the quantities governing the coupling terms, as defined by equation (2.5).

It has been shown (Diver, Ph.D. thesis, 1986) that mode conversion is most likely when the modes do not actually cross but simply approach each other (thus the discriminant, $b^2 - 4ac$, in equation (2.4) has a minimum but does not pass through zero). In order to do this, the eigenvalues were rewritten formally as:

$$\lambda_{1,2} = \alpha \pm \beta.$$

For the case in which α and β both vary linearly with z, the ratio of the asymptotic solutions of equation (2.2) yields a negligible mode conversion coefficient - most of the energy being transmitted. In this case, the eigenvalues cross at the critical point and the extrema of the coupling terms occur at infinity. The corresponding potential must be of the form $f = z_c^2 - z^2$ which is positive for $z \in (-z_c, z_c)$ and negative elsewhere. This is therefore the potential well or underdense potential barrier problem for which appreciable mode conversion is not expected. If, however, α varies linearly with z but $\beta = (b_0 + b_1 z + b_2 z^2)^{1/2}$, there is significant reflection and the coupling reaches its maximum value at points symmetrically placed about the transition point. The degree of coupling is found to depend on the distance between the eigenvalues at their point of closest approach. There therefore appears to be a strong correlation between whether or not the eigenvalues cross and the degree of mode conversion. This agrees with the behaviour anticipated by most authors on physical arguments alone and there is no reason to suppose that such arguments are not true for more general cases of eigenvalue variation.

6. Applications to magnetoionic theory

The general theory of Clemmow and Heading, and in particular the convenient matrix notation, has been adopted by other authors, especially in application to magnetoionic theory. This theory pertains to a *cold plasma*, one in which the random kinetic energy of the particles is taken to be zero, and which contains a magnetic field. (This model will be the focus of our attention in Chapters 5 and 6.) For simplicity, the frequency of wave propagation in magnetoionic theory is often assumed to be high enough that the plasma ions may safely be taken to be fixed. This situation may then be described using Maxwell's equations into which a suitable electric displacement vector $\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}$ (where $\mathbf{P}$ is the electric polarisation) has been substituted. The matrix relating $\mathbf{P}$ to $\mathbf{E}$, the susceptibility matrix, is constructed by substituting the definition $\mathbf{P} = Ne\,\mathbf{r}$ into the equation of motion for an electron moving in a magnetic field and it is composed of elements involving the electron plasma and cyclotron frequencies. The result is a set of four first order coupled linear ordinary differential equations in the components of the electric and magnetic fields; see, for example, chapter 18 of Budden's book on "Waves in the Ionosphere" (1961). If the coefficients of these equations were constant, the determinant could be taken in the usual way in order to derive an equation for the four roots - the *Booker quartic*. Alternatively, the matrix notation of §4 may be used which will also yield the Booker quartic for the homogeneous case.

When the medium is not uniform, the crux of the theory is again the identification of the coefficients of the equations with the appropriate refractive indexes, n, for waves travelling in: "a fictitious homogeneous medium with the properties of the actual ionosphere at each level". Budden recognises that such an approach is limited by the regions of validity of WKBJ theory and limits his analysis accordingly, maintaining close links with the results of WKBJ throughout the relevant parts of his book and emphasising the fact that such approximations must break down at reflection points. In fact he states: " At every place where n is varying, the reflection process is going on, and it is this which prevents the occurrence of a true progressive wave. Except in certain places, however, the process is very weak, so that the WKBJ solutions are very good approximations."

When Budden's method has been taken up by other workers, it often appears that they have missed the crucial importance of the gradient terms since they do not stress the limited domain of validity imposed on the solutions by the relative magnitude of these terms. Fidone and Granata (1971) consider the propagation of waves in a cold plasma confined in a toroidal vessel under the influence of a slightly sheared magnetic field. An initial note of caution should be sounded at this point about the physical implications of having an ambient magnetic field with $\nabla\times\mathbf{B} \neq 0$. In Chapter 6 the full ramifications of this will be investigated, but it is sufficient to mention here that this non-zero term leads to several new terms in the linearised equations of motion and would result in the elements of the susceptibility matrix having different forms. The eigenmodes are significantly altered by the change in the background quantities and it will be shown that the spatially rotating magnetic field of Chapter 6 results in the electric field components having a periodic variation in space dependent on the scalelength of the variation of the magnetic field as well as depending on the homogeneous refractive indices. Because the new terms depend on the magnitude of the gradients, the discrepancy between the complete and approximate solutions should remain small wherever the WKBJ solutions are valid. It is therefore vital to make it clear where these solutions apply. Fidone and Granata calculate the degree of mode coupling from the ratio of the electric field components on either side of the conversion region but these solutions for the components are derived from an iterative method which is only valid far from the coupling region and it is not clear that conversion coefficients may be calculated entirely from these functions without reference to the known behaviour of the coefficients in the interaction area. It is, after all, the specific dependence of these quantities and their gradients which causes the mode conversion.

An early paper employing Budden's approach to magnetoionic theory was that by Frisch (1964) which examined the coupling of MHD waves in stratified media with particular reference to the heating of the solar corona. The original idea came from, amongst others, Osterbrock (1961) who suggested that sound waves excited in the convection zone could propagate through the photosphere and undergo mode conversion to Alfvén waves in the high chromosphere or corona, leading to energy deposition at these levels. Such a change in wave types was necessary since Alfvén waves originating in the convection zone would be absorbed completely before attaining chromospheric or coronal heights and such a direct mechanism could not therefore explain the anomalously high coronal temperature.

In Frisch's paper, the magnetic field is allowed to vary in both magnitude and direction but, again, the equilibrium current required to balance the effect of the changing magnetic field is omitted. Because the background to the WKBJ solutions and the inherent approximations are not stated explicitly, the implication seems to be that the solutions inferred by the matrices are applicable everywhere. Frisch's main aim is to calculate the elements of the coupling matrix which is taken to indicate the strength of the interaction between "near neighbours" - modes labelled by their homogeneous wavenumber, and these are used as a measure of the degree of mode coupling. He notes that the elements of the coupling matrix will be largest for large parameter gradients and modes which are close together - this result was also deduced in §4. The elements of this coupling matrix, which is equivalent to $\mathbf{A}^{-1}\mathbf{A}'$, are used as an absolute measure of the coupling present. This is not the same as calculating how much of a second mode will be excited at a coupling point by a known input. The latter process generally involves solving the full ode and analytically continuing the asymptotic solutions through the transition region in order to calculate coefficients of transmission etc. It was shown in §4 that the size of these elements alone is not enough to determine the degree of mode conversion which depends strongly on the exact behaviour of the modes in the interaction region.

A series of papers (Melrose (1974a&b), Melrose (1977a&b), Melrose and Simpson (1977), Melrose (1980)) on the subject of mode coupling in the solar corona, follows closely the work of Frisch. Although Melrose describes the parameters as "slowly varying", this is not quantified in terms of the ratio of the magnitude of the parameter gradients to the parameters themselves and it remains unclear that the work is being carried out under the limitations of WKBJ theory. Melrose identifies the modes of the system throughout by the roots of the homogeneous dispersion relation which again vary because of the spatial dependence of the physical parameters (predominantly the magnetic field). The results presented regarding the extent of mode coupling are again derived from the magnitude of the elements of the coupling matrix, and in addition, the author identifies possible mode conversion points as occurring wherever roots of the dispersion relation are equal. It is precisely in the neighbourhood of the coupling points that the approximation of employing solutions from homogeneous magnetoionic theory breaks down and more accurate solutions of the differential equation are required. It was also indicated in §4 that mode conversion is probably more likely to be non-negligible for eigenvalues which approach each other without crossing than for ones which are equal at some point and that the degree of mode conversion is then tied to the minimum separation of the modes. Melrose calculates the wave properties around the coupling point from the known properties for a homogeneous medium using a perturbation approach. This "coupling approximation" assumes on one hand that the homogeneous wavenumbers provide adequate solutions in this region while also supposing that coupling is large, and therefore that in this neighbourhood the gradient terms, which are absent in the calculation of the k's, are much larger here than they are elsewhere.

7. Summary

In this chapter, we have discussed the propagation of waves in an inhomogeneous medium with particular attention being paid to the possibility of the modes exchanging their character. In order to do this, it was necessary to find a method of identifying the possible modes of propagation since the progressive, sinusoidal waves which occur in homogeneous media are no longer appropriate in the non-uniform case because of the presence of parameter gradients. For slowly-varying media (which we defined precisely), the WKBJ solutions provide a good approximation to the correct solution but these become invalid near points where the wave potential becomes small or the gradients become too large - these are points where mode conversion may take place. In the case of two normal modes, mode conversion manifests itself as reflection and is thus a long-recognised phenomenon. The flaws in using local dispersion relation approaches were discussed along with the major difficulties involved in calculating mode conversion coefficients. Finally, applications of the eigenvalue method to magnetoionic theory, MHD and solar physics were reviewed and discussed.

Having dismissed much of the existing mode conversion literature, we must attempt to replace it with a more acceptable alternative. We will first compare and contrast the dispersion relation and eigenvalue approaches to atmospheric wave propagation (Chapter 4) as a straightforward example of waves in an inhomogeneous medium. This will be followed, in Chapter 6, by a complete, analytic solution of the problem of waves propagating in a cold plasma with a spatially rotating magnetic field.

Chapter 3 - Derivation of Model Equations

1. Introduction

The fluid model of a plasma has its roots in Kinetic Theory and differs from standard fluid mechanics only in that it considers the effects of electromagnetic forces. By deriving from kinetic theory the familiar equations of fluid mechanics (Landau and Lifshitz, 1959) as they apply to a conducting fluid, we aim to demonstrate why the "cold plasma" to be discussed in Chapters 5 and 6 may be adequately described by virtually the same equations as used to describe the neutral atmosphere of Chapter 4. In the process, we will also be able to highlight the approximations which are necessary in order to make the cold plasma equations applicable. Because these equations form the backbone of our description of wave propagation, understanding their origins is vital and so we follow their derivation in detail.

Kinetic Theory is a statistical method of describing a plasma. It is by far the most rigorous theoretical description of the plasma state and, because it is more comprehensive than its derivative orbit and fluid theories, kinetic theory is in general more complicated, frequently requiring numerical methods of solution. Once the distribution of probabilities for finding a specific type of particle (electron or ion) in 6-dimensional coordinate-velocity space has been obtained, all relevant properties of the plasma may be determined. In order to make the problem more tractable, the equations describing the plasma's microscopic properties are often combined in such a way that they describe measurable, fluid-like, macroscopic quantities such as density and pressure.

2. The general transport equation

The hydrodynamic equations describing the evolution of such macroscopic fluid quantities as density (ρ), velocity ($\mathbf{u}$) and pressure (p) may be derived in one of two ways. Either one may use plausibility arguments to extend the fluid equations by analogy to a conducting fluid to produce conservation equations or one may take the "moments" of the Boltzmann equation, yielding the same set of evolution equations. The Boltzmann equation itself may be derived from either the Liouville equation for the evolution of an N-particle distribution function (Boyd and Sanderson, 1969) or using a more heuristic argument (ter Haar, 1954; Chapman and

Cowling, 1960).

We derive the fluid equations from the Boltzmann equation using the method of moments, which simply involves multiplying by velocity to some power and averaging over velocity space. In theory, this method would generate an infinite system of coupled equations because each moment introduces an additional variable so that the number of the variables is always greater than the number of equations. The zero-order moment leads to the continuity equation, containing both the scalar density and vector velocity, the first- order equation of motion contains the velocity and the pressure dyad and so on. To obtain an equation for the evolution of each new quantity would require taking the next moment and so the process would continue ad infinitum unless we truncate the hierarchy at some stage using physical arguments - either by neglecting the moments above a certain velocity power, or by assuming an equation of state relating p to ρ. The method of closure determines a model for the plasma which, in turn, determines which plasma properties can be studied, since the approximations used eliminate certain features of plasma behaviour. Thus the area of validity of our set of equations is necessarily limited to the regions in which these approximations are valid.

The Boltzmann equation, which describes the evolution of a one-body distribution function, f_s, for particle species s is:

$$\frac{\partial f_s}{\partial t} + \mathbf{v}\cdot\nabla f_s + \mathbf{a}\cdot\nabla_{\mathrm{v}} f_s = \left(\frac{\partial f_s}{\partial t}\right)_{\mathrm{coll}} \tag{3.1}$$

where

1) $f_s(\mathbf{r},\mathbf{v},t)$ is the distribution function which gives the probability of finding particles of type s within the 6-dimensional volume element $\mathbf{dr}\,\mathbf{dv}$ centred at the point $(\mathbf{r},\mathbf{v})$ in coordinate and velocity space.

2) ∇_{v} is the gradient acting in velocity space .

3) $\mathbf{a}$ is the acceleration term arising from any forces acting. The external forces which might be considered are gravity ($\mathbf{F} = m_s\mathbf{g}$, Chapter 4), and those due to an applied electric field ($\mathbf{F} = q_s\mathbf{E}$) and a magnetic field ($\mathbf{F} = \mathbf{v}\times\mathbf{B}$, Chapters 5 & 6). The last of these differs from the other two by being velocity-dependent and must therefore be treated carefully when taking moments. This acceleration term also depends on the self-consistent fields which are generated by charged particles moving and because these fields depend on the distribution function, equation (3.1) is non- linear. The remaining interactions which the particles suffer are rapid fluctuations on the microscopic scale and these are collected together in the "collision" term on the right hand side of equation (3.1) and treated separately. When this term is neglected, equation

(3.1) is known as the collisionless Boltzmann, or Vlasov, equation.

Consider a quantity, ϕ , representing any physical property of the particles (which may, in general, be a function of velocity). The average value of $\phi(\mathbf{v})$ with respect to velocity space of the particles of type s (denoted by $< >_s$) is defined to be

$$<\phi>_s \stackrel{\partial}{=} \frac{1}{n_s(\mathbf{r},t)} \int_{\mathbf{v}} \phi(\mathbf{r},\mathbf{v},t) f_s(\mathbf{r},\mathbf{v},t) d\mathbf{v}$$

where $n_s(\mathbf{r},t) = \int_{\mathbf{v}} f_s(\mathbf{r},\mathbf{v},t)\, d\mathbf{v}$ is the number density of particles of type s in spatial volume element $d\mathbf{r}$. Thus, generating the corresponding evolution equation for the macroscopic quantity $<\phi>_s$ requires a similar averaging over velocity space to be performed on the whole of equation (3.1). In the general case, multiplying each term in turn by ϕ gives:

First term

$$\int_{\mathbf{v}} \phi(\mathbf{v}) \frac{\partial f_s}{\partial t} d\mathbf{v} = \frac{\partial}{\partial t} \int_{\mathbf{v}} \phi f_s d\mathbf{v} = \frac{\partial}{\partial t} n_s <\phi>_s ,$$

since $\mathbf{v}$ and t are independent variables and ϕ is not time-dependent.

Second term (similarly, because $\mathbf{r}$ and $\mathbf{v}$ are independent variables)

$$\int_{\mathbf{v}} \phi(\mathbf{v}) \mathbf{v} \cdot \nabla f_s d\mathbf{v} = \nabla \cdot \int_{\mathbf{v}} \phi \mathbf{v} f_s d\mathbf{v} = \nabla \cdot (n_s <\phi \mathbf{v}>_s),$$

Third term (for forces which are velocity-independent)

$$\int_{\mathbf{v}} \phi(\mathbf{v}) \mathbf{a} \cdot \nabla_{\mathbf{v}} f_s d\mathbf{v} = \mathbf{a} \cdot \int_{\mathbf{v}} (\nabla_{\mathbf{v}} (\phi f_s) - f_s \nabla_{\mathbf{v}} \phi) d\mathbf{v} \tag{3.2}$$

$$= -n_s <\mathbf{a} \cdot \nabla_{\mathbf{v}} \phi>_s ,$$

where the first term of equation (3.2) vanishes since we assume that $f_s \to 0$ sufficiently rapidly as $|\mathbf{v}| \to \infty$ so that $\lim_{|\mathbf{v}| \to \infty} (\phi f_s) = 0$ for all functions $\phi(\mathbf{v})$. The only velocity- dependent field which will be of interest to us will be magnetic in origin and, since the ith component of $\mathbf{v} \times \mathbf{B}$ does not contain v_i, the above procedure may be repeated.

Adding these terms gives the general transport equation which describes the transport of such macroscopic quantities as mass, momentum, energy etc., depending on the choice of ϕ:

$$\frac{\partial}{\partial t}(n_s <\phi>_s) + \nabla \cdot (n_s <\phi \mathbf{v}>_s) - n_s <\mathbf{a} \cdot \nabla_{\mathbf{v}} \phi>_s = \left(\frac{\partial}{\partial t} (n_s <\phi>_s) \right)_{coll} \tag{3.3}$$

3. Specific moments of the Boltzmann equation

The functions to be substituted for $\phi(\mathbf{v})$ are chosen to ensure that the equations generated relate to physical observables. Henceforth, we will be concerned with a fully ionised plasma consisting of ions and electrons.

3.1. Continuity equation

We take the zero-order moment, setting $\phi(\mathbf{v}) = 1$, so that the terms of equation (3.3) become (in order)

$$\frac{\partial}{\partial t}(n_s \langle\phi\rangle_s) = \frac{\partial}{\partial t} n_s ,$$

$$\nabla\cdot(n_s \langle\phi \mathbf{v}\rangle_s) = \nabla\cdot(n_s \mathbf{u}_s)$$

$$-n_s \mathbf{a}\cdot\nabla_v \phi = 0$$

where we have defined $\mathbf{u}_s$ to be the average velocity of a particle of species s. Substituting these in equation (3.3) gives the density conservation equation, or continuity equation,

$$\frac{\partial n_s}{\partial t} + \nabla\cdot(n_s \mathbf{u}_s) = S_s \qquad (3.4)$$

where $S_s = (\frac{\partial n_s}{\partial t})_{coll}$, represents the rate per unit volume at which particles of type s are produced as a result of collisions. Physically, this term is composed of processes of particle creation or destruction such as ionisation and recombination which we will henceforth assume to be negligible.

3.2. Momentum conservation equation

In this section, we will require several new definitions. The first is the concept of mass density which is defined in terms of the number density to be $\rho_s \equiv n_s m_s$. (The equation of mass conservation may easily be obtained by multiplying equation (3.4) by the mass of a particle of type s.)

We split the particle velocity into two separate contributions - one due to the flow of the fluid, $\mathbf{u}_s$, and the other denoting the random particle motions with respect to this flow due to their thermal energy, $\mathbf{w}_s$, so that $\mathbf{v} = \mathbf{u}_s + \mathbf{w}_s$. We note for later use that $\langle\mathbf{w}_s\rangle = 0$ because of the random nature of the thermal motions.

We define the pressure on an imaginary surface element moving with the average fluid velocity to be the rate of transport of molecular momentum per unit area due to random particle motions, so that the pressure dyad, $\mathbf{p}_s \overset{\partial}{=} \rho_s \langle\mathbf{w}_s \mathbf{w}_s\rangle$. The diagonal components of $\mathbf{p}_s$ are analogous to hydrostatic pressure acting normal to some conceptual surface in the plasma while the off-diagonal terms are pressures due to viscosity and shear stresses acting tangential to these surfaces. Another important macroscopic variable is the mean hydrostatic or scalar pressure which is defined to be one third of the trace of the pressure tensor and is therefore given by $p_s = 1/3\rho_s <w_s^2>$. Finally, the absolute temperature of the particles of type s is a measure of the mean kinetic energy of the random particle motion and is related to the scalar pressure by the equation of state for an ideal gas, $p_s = n_s k_B T_s$, where k_B is Boltzmann's constant.

For the first-order moment, we take $\phi(\mathbf{v}) = m_s \mathbf{v}$ so that the terms of equation (3.3) then become:

$$\frac{\partial}{\partial t}(n_s <m_s \mathbf{v}>_s) = \frac{\partial}{\partial t}(\rho_s \mathbf{u}_s)$$

$$= \mathbf{u}_s \frac{\partial \rho_s}{\partial t} + \rho_s \frac{\partial \mathbf{u}_s}{\partial t},$$

$$\nabla \cdot (n_s <m_s \mathbf{v}\mathbf{v}>_s) = \nabla \cdot (\rho_s (\mathbf{u}_s \mathbf{u}_s + 2\mathbf{u}_s <\mathbf{w}_s> + <\mathbf{w}_s \mathbf{w}_s>))$$

$$= \rho_s (\mathbf{u}_s \cdot \nabla)\mathbf{u}_s + \mathbf{u}_s \nabla \cdot (\rho_s \mathbf{u}_s) + \nabla \cdot \mathbf{p}_s,$$

$$-n_s <\mathbf{F} \cdot \nabla_\mathbf{v} \mathbf{v}>_s = -n_s <(F_x \frac{\partial}{\partial v_x} + F_y \frac{\partial}{\partial v_y} + F_z \frac{\partial}{\partial v_z})\mathbf{v}>_s$$

$$= -n_s <(F_x \hat{\mathbf{x}} + F_y \hat{\mathbf{y}} + F_z \hat{\mathbf{z}})>_s$$

$$= -n_s <\mathbf{F}>_s .$$

We combine the constituent terms of equation (3.3) and substitute from the continuity equation, where we again assume that no particles are being created or destroyed, forming the momentum conservation equation:

$$\rho_s (\frac{\partial \mathbf{u}_s}{\partial t} + (\mathbf{u}_s \cdot \nabla)\mathbf{u}_s) + \nabla \cdot \mathbf{p}_s - n_s <\mathbf{F}>_s = \mathbf{A}_s \tag{3.5}$$

where $\mathbf{A}_s = (\frac{\partial(\rho_s \mathbf{u}_s)}{\partial t})_{coll}$ takes account of transfer of energy between different particle species because a collision between particles of the same type conserves momentum amongst those particles. Since the collision term represents force per unit volume exerted on particles of type s due to collisions with particles of type α, it seems plausible* to make it dependent on the difference in velocities between the two particles: $\mathbf{A}_s = -\rho_s \sum_s \nu_{s\alpha}(\mathbf{u}_s - \mathbf{u}_\alpha)$ where the constants of proportionality, $\nu_{s\alpha}$

* For a varying temperature profile a ∇T term would also contribute to the frictional force.

and $\nu_{\alpha s}$ are the *collision frequencies* for collisions between particles of the specified species. The total momentum in the system must be conserved and so the relation $\rho_i \nu_{ie} = \rho_e \nu_{ei}$ holds where the subscripts i and e refer to ions and electrons respectively.

It is convenient to introduce the concept of the *advective derivative* in relation to equation (3.5)

$$\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + \mathbf{u}_s \cdot \nabla .$$

This derivative describes the rate of change of a quantity in the reference frame of a fluid element travelling with the fluid at the average fluid velocity. The first contribution represents temporal variation in the quantity viewed from a fixed point and the second allows for the motion of the fluid element in this time - it has moved (because of the flow) to a region where the quantity is different. This second factor will obviously be zero when either the physical quantity under consideration does not vary in space or when the fluid is at rest. Using this notation and assuming that the only forces acting are electromagnetic, we may rewrite equation (3.5) as:

$$\rho_s \frac{\mathrm{D}\mathbf{u}_s}{\mathrm{D}t} = n_s q_s (\mathbf{E} + \mathbf{u}_s \times \mathbf{B}) - \nabla \cdot \mathbf{p}_s + \mathbf{A}_s .$$

This equation expresses the fact that the time rate of change of the mean momentum in each fluid element is due to the external forces applied to the fluid, to the shear and pressure forces of the fluid itself and to the internal forces associated with the collisional interactions.

3.3. Energy equation

The new definition which we must introduce here before proceeding is that for the heat flow triad, $\mathbf{q}_s \overset{\partial}{=} 1/2 \rho_s \langle w_s^2 \mathbf{w}_s \rangle$, the random thermal energy flux due to the particles' random thermal motions.

Taking $\phi(\mathbf{v}) = 1/2 m_s \mathrm{v}^2$ in our second-order moment, the first two terms of the general transport equation (3.3) become:

$$\frac{\partial}{\partial t}(n_s \langle 1/2 m_s \mathrm{v}^2 \rangle_s) = \frac{\partial}{\partial t}(1/2\, \rho_s u_s^2 + 1/2\, \rho_s \langle w_s^2 \rangle) = \frac{\partial}{\partial t}(1/2\, \rho_s u_s^2 + 3/2 p_s),$$

$$\nabla \cdot (n_s \langle 1/2 m_s \mathrm{v}^2 \mathbf{v} \rangle_s) = \nabla \cdot (1/2\, \rho_s \langle (\mathbf{v} \cdot \mathbf{v}) \mathbf{v} \rangle_s)$$

$$= \nabla \cdot (1/2\, \rho_s (u_s^2 \mathbf{u}_s + \langle w_s^2 \rangle \mathbf{u}_s + 2\mathbf{u}_s \cdot \langle \mathbf{w}_s \mathbf{w}_s \rangle + \langle w_s^2 \mathbf{w}_s \rangle))$$

$$= \nabla \cdot (1/2\, \rho_s u_s^2 \mathbf{u}_s) + 3/2 p_s \nabla \cdot \mathbf{u}_s + 3/2 (\mathbf{u}_s \cdot \nabla) p_s + \nabla \cdot (\mathbf{p}_s \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{q}_s ,$$

where we have used the definition of the scalar pressure from the previous section and a vector identity to separate out the constituent parts of the term $\nabla\cdot(3/2\,p_s\mathbf{u}_s)$. Employing the advective derivative enables us to write :

$$\frac{\mathrm{D}}{\mathrm{D}t}(\frac{3}{2}p_s)+\frac{3}{2}p_s\nabla\cdot\mathbf{u}_s+\frac{\partial}{\partial t}(\frac{1}{2}\rho_s u_s^2)+\nabla\cdot(\frac{1}{2}\rho_s u_s^2\mathbf{u}_s)+\nabla\cdot(\mathbf{p}_s\cdot\mathbf{u}_s)$$

$$+\nabla\cdot\mathbf{q}_s-n_s<\mathbf{F}\cdot\mathbf{v}>_s\;=\mathbf{M}_s, \qquad (3.6)$$

where $\mathbf{M}_s\;=(\frac{\partial}{\partial t}(1/2\,\rho_s<v^2>_s)_{coll})$ is the rate of energy density change due to collisions and, like $\mathbf{A}_s$, contains nonzero contributions from energy transfer between species.

The third and fourth terms of equation (3.6) may be combined in the form:

$$\tfrac{1}{2}u_s^2(\frac{\partial\rho_s}{\partial t}+\nabla\cdot(\rho_s\mathbf{u}_s))+\rho_s\mathbf{u}_s\cdot\frac{\mathrm{D}\mathbf{u}_s}{\mathrm{D}t},$$

the first member of which is clearly zero (by comparison with the continuity equation) since we have chosen to ignore particle creation and destruction processes (i.e. $S_s\;=0$) and the second of which may be substituted for using equation (3.5). The two force terms now sum to

$$n_s\mathbf{u}_s\cdot<\mathbf{F}>_s-n_s<\mathbf{F}\cdot\mathbf{v}>_s\;=-n_s<\mathbf{F}\cdot\mathbf{w}_s>,$$

which is automatically zero for velocity-independent forces and also for the Lorentz force since it acts orthogonally to the velocity. Using the identity $\nabla\cdot(p_s\cdot\mathbf{u}_s)-\mathbf{u}_s\cdot(\nabla\cdot p_s)=(p_s\cdot\nabla)\cdot\mathbf{u}_s$, we may state the equation of energy conservation in the final form:

$$\frac{\mathrm{D}}{\mathrm{D}t}(\frac{3}{2}p_s)+\frac{3}{2}p_s\nabla\cdot\mathbf{u}_s+(\mathbf{p}_s\cdot\nabla)\cdot\mathbf{u}_s+\nabla\cdot\mathbf{q}_s\;=\mathbf{M}_s-\mathbf{u}_s\cdot\mathbf{A}_s. \qquad (3.7)$$

In physical terms, the first member of equation (3.7) is the total rate of change of the particle thermal energy density moving with a fluid element. The second is the change in thermal energy density due to particles entering the fluid element with the average fluid velocity, $\mathbf{u}_s$. The third is related to the work done by the kinetic pressure inside the fluid while the fourth allows for heat flux. The terms on the right hand side of equation (3.7) represent changes in the thermal energy density caused by external forces and collisions respectively.

4. Electromagnetic variables

There still remain the electromagnetic variables **E** and **B** for which evolution equations must also be provided. We use the last two Maxwell's equations:

Faraday's Equation

$$\nabla\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{3.8}$$

and the generalisation of Ampère's equation

$$\nabla\times\mathbf{B} = \mu_0(\mathbf{J}+\varepsilon_0\frac{\partial \mathbf{E}}{\partial t}), \tag{3.9}$$

where **J** is the current which is caused by the motion of charged particles.

(Note that the equation $\nabla\cdot\mathbf{B} = 0$ may be viewed as an initial condition on equation (3.8) and need not be stated explicitly itself although the fact that it always holds is understood.)

5. Closure

As predicted, the set of equations which has now been generated contains more unknowns than equations, i.e. the system is not yet closed. At this stage, the closure assumptions which are made define the model under consideration. The more restrictive these assumptions, the more limited will be the physics which the model can describe. Already the assumption that the plasma behaves as a conducting fluid has ruled out the possibility of describing microscopic particle motions.

5.1. The cold plasma model

This represents the simplest method of closing the set of fluid equations. In hydrodynamics, the persistence of fluid elements is ensured by collisions with neighbouring particles which tend to keep particles in a particular fluid element and so the size of the element must be much larger than the mean free path, $d\mathbf{r} \gg \lambda_c$. Also, by definition, a fluid element must be much smaller then the characteristic length over which the hydrodynamic variables change, $d\mathbf{r} \ll \lambda$. Combining these results gives a condition for a meaningful description of a fluid element :

$$\lambda \gg \lambda_c .$$

In a plasma, however, this condition often proves far too restrictive. Also, we wish to use the fluid equations to describe a plasma in which the effect of collisions is negligible and the coherence of the fluid elements must be enforced by another means. In fact, the electromagnetic forces acting between the plasma particles are sufficient to maintain the fluid behaviour even in the absence of collisions as long as

$$|\mathbf{u}| \gg |\mathbf{w}|.$$

Small random velocities correspond to pressure and temperature also being negligible, by their definitions. We therefore set $\mathbf{q}_s = 0$ and $\mathbf{p}_s = 0$ so that only equations (3.4) and (3.5) are required to describe the evolution of the hydrodynamic variables in the cold plasma model. The set of cold plasma equations is completed by including equations (3.8) and (3.9) for the electric and magnetic fields plus an equation for the current, $\mathbf{J} = \sum_s n_s q_s \mathbf{u}_s$. These equations constitute the description of the cold plasma model and provide us with the set of 15 equations in 15 unknowns (including both species) which we will use extensively in Chapters 5 and 6 and which are listed together, for convenience, at the beginning of Chapter 6.

A description such as this in which the ions and electrons are regarded separately is known as a two-fluid model and contrasts with the usual MHD equations in which the ions and electrons are regarded as constituting a single fluid.

5.2. The single fluid (MHD) model

This model is opposite to the cold plasma model in that collisions are assumed to be the dominant process. Under the effect of multiple collisions, the particle distribution function relaxes to a Maxwellian in the order of a few collision times. As a result, the elements of the heat flux tensor become small, the pressure tensor takes on a particularly simple form and the pressure and temperature of the ions and electrons become approximately equal since the thermal conductivity and viscosity are related. These basic assumptions provide our new method of closing the system:

1)

$$q_{ijk} = 0 \quad \text{for all } i, j, k$$

2)

$$\mathbf{p}_s = \begin{bmatrix} p_s & 0 & 0 \\ 0 & p_s & 0 \\ 0 & 0 & p_s \end{bmatrix},$$

where p_s is the scalar pressure and the term $\nabla \cdot \mathbf{p}_s$ becomes ∇p_s throughout.

To make our set of equations smaller, we average the equations over all species and consider velocities to be relative to a weighted mean flow (using the subscripts e and i for electrons and ions respectively):

$$\mathbf{u} \equiv \frac{m_e n_e \mathbf{u}_e + m_i n_i \mathbf{u}_i}{m_e n_e + m_i n_i}.$$

Constructing a single fluid description of the plasma from the two-fluid equations involves a number of definitions for similarly averaged quantities. Assuming that there is only one type of ion for simplicity, we define :

$$\rho \equiv m_e n_e + m_i n_i \quad , \quad q \equiv e(n_i - n_e) \quad , \quad \mathbf{J} \equiv e(n_i \mathbf{u}_i - n_e \mathbf{u}_e),$$

where e is the electronic charge. Overall, the plasma must be neutral and so $n_e = n_i = n$. The electron and ion velocities must be approximately equal (since $m_e \ll m_i$) so that $\mathbf{u}_e \approx \mathbf{u}_i \approx \mathbf{u}$ and therefore, by definition, $p_e \approx p_i \approx P/2$ where P is the total scalar pressure .

The averaging procedure removes any difficulty with assigning values to the collision terms in the momentum and energy conservation equations for the single fluid because momentum lost by electrons will be gained by the ions and so, for the system as a whole, both momentum and energy are conserved. Having averaged over the species, an equation for the evolution of charge density can easily be derived from the continuity equation (3.4) by multiplying each equation by the charge on the appropriate species then adding the results to get :

$$\frac{\partial q}{\partial t} + \nabla \cdot \mathbf{J} = 0.$$

An equation for the evolution of the current is now required to complete the set - it may be constructed in a similar way to the charge density equation by multiplying the momentum conservation equations by q_s / m_s and adding. Here, however, the collision terms are no longer equal and opposite but add to give:

$$\frac{e}{m_i}\mathbf{A}_i - \frac{e}{m_e}\mathbf{A}_e = e\,\mathbf{A}_i\left(\frac{1}{m_i} - \frac{1}{m_e}\right)$$

$$= -\frac{e}{m_e} m_e n_e \nu_{ei}(\mathbf{u}_i - \mathbf{u}_e) \approx -\nu_{ei}\mathbf{J},$$

where we have used the approximations stated above. The complete equation then becomes, on neglecting terms of order J^2 and u^2, the generalised Ohm's Law:

$$\frac{m_e}{ne^2}\frac{\partial \mathbf{J}}{\partial t} = \mathbf{E} + \mathbf{u}\times\mathbf{B} - \frac{1}{ne}(\mathbf{J}\times\mathbf{B}) + \frac{\nabla P}{2ne} - \frac{\mathbf{J}}{\sigma} \qquad (3.10)$$

where $\sigma = \dfrac{ne^2}{m_e \nu_{ei}}$ represents the electrical conductivity.

Using dimensional arguments, we can demonstrate that several terms of the MHD equations make negligible contributions and so may be dropped. From the first of Maxwell's curl equations, equation (3.8), we find that the ratio $E/B \sim \omega L$ where ω^{-1} and L are characteristic time and length scales over which the the fields change appreciably. To give the field and flow effects equal status, we must have $u \sim \omega L$ and so, since the plasma is non-relativistic, $\omega L/c \ll 1$. As a result, we find that the displacement current term in equation (3.9) may be neglected since

$$\mu_0 \varepsilon_0 \frac{\partial E}{\partial t} / \nabla \times B \approx \frac{\frac{E}{B}\omega L}{c^2} \sim \left(\frac{\omega L}{c}\right)^2 \ll 1.$$

We now extend this simple dimensional analysis to evaluate the relative importance of each term of equation (3.10). Normalising with respect to the electric field, we find the terms of (3.10) are in the ratios:

$$\frac{\omega^2}{\omega_{pe}^2}\frac{c^2}{u^2}:1:1:\frac{\omega}{\omega_{pe}}\frac{\Omega_e}{\omega_{pe}}\frac{c^2}{u^2}:\frac{\omega}{\Omega_i}\frac{c_s^2}{u^2}:\frac{\omega}{\omega_{pe}}\frac{\nu_{ei}}{\omega_{pe}}\frac{c^2}{u^2}$$

where we have expressed these ratios in terms of the ion and electron cyclotron frequencies, $\Omega_i = eB/m_i$, $\Omega_e = eB/m_e$ and the electron plasma frequency, $\omega_{pe} = ne^2/\varepsilon_0 m_e$, the electron-ion collision frequency and the sound speed, $c_s \sim (P/\rho)^{1/2}$. It is thus clear that, depending on the frequency regime under consideration, a different subset of the terms of equation (3.10) is required for an accurate description. When the frequencies of interest are sufficiently low, the $\partial \mathbf{J}/\partial t$, $\mathbf{J}\times\mathbf{B}$ and ∇P terms may be neglected and the appropriate form of Ohm's Law is then simply:

$$\mathbf{J} = \sigma(\mathbf{E} + \mathbf{u}\times\mathbf{B}). \tag{3.11}$$

It is often assumed that the conductivity is perfect so that, by equation (3.11) , $\mathbf{E} + \mathbf{u}\times\mathbf{B} = 0$. To combine the ion and electron versions of the energy equation into the MHD approximation, it is best to return to equation (3.6), bearing in mind that the velocities are now relative to $\mathbf{u}$ and not $\mathbf{u}_s$. A little algebra yields:

$$\frac{3}{2}\frac{\mathrm{D}P}{\mathrm{D}t} + \frac{5}{2}P\,\nabla\cdot\mathbf{u} = (\mathbf{J} - q\,\mathbf{u})\cdot(\mathbf{E} + \mathbf{u}\times\mathbf{B}). \tag{3.12}$$

In the limit of perfect conductivity, the right hand side of equation (3.12) vanishes and, substituting for $\nabla\cdot\mathbf{u}$ from the continuity equation, (3.4), we find:

$$\frac{\mathrm{D}}{\mathrm{D}t}\left(\frac{3}{2}P\right) - \left(\frac{5}{2}\frac{P}{\rho}\right)\frac{\mathrm{D}\rho}{\mathrm{D}t} = 0,$$

$$\frac{D}{Dt}(P\rho^{-5/3}) = 0,$$

$$P\rho^{-5/3} = constant,$$

which is the adiabatic equation of state for a gas which has the ratio of its specific heat capacities, at constant pressure and constant volume respectively , equal to 5/3. (This only holds when viscosity, thermal conductivity and the elements of the heat flux tensor are all negligible.) Also when $\sigma \to \infty$, we may construct an alternative equation to (3.8) which will prove useful in the succeeding chapter. From Faraday's Law:

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla\times\mathbf{E} = \nabla\times(\mathbf{u}\times\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{u} - (\nabla\cdot\mathbf{u})\mathbf{B} \tag{3.13}$$

where we have used a vector identity and assumed that the magnetic field is homogeneous in space.

When the kind of dimensional analysis shown above is applied to the remainder of the MHD equations, it is found that the charge density may be omitted altogether since,

in the momentum equation (3.5): $qE/JB \sim E^2/B^2c^2 \sim u^2/c^2$,

in the energy equation (3.12) (for the case of nonperfect conductivity): $qu/J \sim Eu/Bc^2 \sim u^2/c^2$

where we have used Poisson's equation ($\nabla\cdot\mathbf{E} = q/\varepsilon_0$) to relate the total charge density and the electric field.

Thus the full set of single fluid MHD approximations for a perfectly conducting plasma (the ideal hydromagnetic equations) may be written:

$$\frac{D\rho}{Dt} = -\rho\nabla\cdot\mathbf{u},$$

$$\rho\frac{D\mathbf{u}}{Dt} = -\nabla P + \mathbf{J}\times\mathbf{B},$$

$$\frac{D}{Dt}(P\rho^{-5/3}) = 0,$$

$$\mathbf{E} + \mathbf{u}\times\mathbf{B} = 0,$$

$$\nabla\times\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$

$$\nabla\times\mathbf{B} = \mu_0\mathbf{J}.$$

6. Summary

We have demonstrated how a plasma may be considered as a conducting fluid under diverse circumstances - when collisions are negligible and when they are dominant. The exact form of the equations differs in the two cases but the philosophy remains the same - the motion of individual plasma particles need not be calculated in order to find certain of the macroscopic plasma parameters and a full kinetic theoretical approach is not required. The cold plasma model is the more restrictive of the but has been used extensively to great effect in the study of plasma waves - magnetoionic theory. Because the random particle motions are ignored, this model cannot detect any finite temperature effects (e.g. acoustic waves) and when these features become important (for very small wave phase velocity, i.e. $c_s \approx v_{ph}$), the cold plasma model must be discarded in favour of an MHD treatment. The two treatments merge at low frequencies where the cold plasma reproduces the results of "pressureless MHD".

Chapter 4 - Atmospheric Acoustic - Gravity Waves

1. Introduction

In this chapter, we present a critique of the work which has been done over the last twenty years or so on atmospheric waves and apply the eigenvalue analysis of Chapter 2 to a neutral atmosphere with vertical temperature stratification. Although an analytic solution of the problem is seldom possible, the eigenvalues may always be calculated and used as a comparison for testing numerical and approximate dispersion relation approaches alike.

A compressible atmosphere in the presence of a gravitational field can support a type of wave motion which at high frequencies becomes the familiar acoustic wave but at very low frequencies tends to a gravity wave (the characteristics of which will be described later). These atmospheric waves are therefore collectively known as *acoustic-gravity* waves. Much early work in this field was prompted by attempts to understand the atmospheric response to great natural disasters like the eruption of Krakatoa in 1883 (Pekeris, 1939), the Siberian meteorite of 1908 (Pekeris, 1948; Scorer, 1950) and powerful earthquakes such as the one in Alaska in 1964 (Mikumo, 1968). Barometric recordings of the Krakatoan eruption also stimulated Lamb to include a section on atmospheric waves in his famous text on hydrodynamics (1945). A renewed interest in such waves arose with the advent of nuclear explosions and the possibility of determining properties of the source from wave characteristics observed at some distance away. With the discovery of large scale oscillations on the Sun, the interest in the propagation of gravity waves spread to astronomy. Many authors have concentrated on the analysis of isothermal model atmospheres or on extensions of the results of this model to cases of non-zero temperature gradient. The latter technique is fundamentally the *local dispersion relation* approach described in Chapter 2 where its shortcomings were highlighted. The emphasis has been on *ducted modes* where some particular atmospheric variable (often temperature) varies with height, providing a waveguide in which the gravity wave can travel. Such *trapping* schemes have been used to explain the solar 5-minute oscillation and terrestrial observations of gravity waves at large horizontal distances away from their point of origin.

Some efforts have been made on the question of mode conversion but these have been severely limited by the small number of functions, $T(z)$, for which a full analytical solution may be obtained. The slant towards ducting models has also hampered this area of study and we hope that it may be revived by the application of the techniques being developed in other disciplines where the problem of wave propagation in inhomogeneous media is of paramount and urgent importance - especially plasma physics.

2. Atmospheric structure

2.1. The Earth's atmosphere

The Earth's atmosphere is naturally inhomogeneous over many different scalelengths. Vertically, it is loosely divided into regions with similar characteristics extending over several kilometres (CIRA, 1965, 1972) with the neutral atmosphere being subdivided according to large scale temperature structures into four layers (Figure (4.1)). In the *troposphere*, the lowest region, the temperature decreases with altitude at a rate of about $-5°\mathrm{K\,km^{-1}}$. The troposphere terminates at the tropopause around 7-17 km depending on the latitude and the season and above this lies the upper atmosphere. The first upper atmospheric layer, the *stratosphere*, possesses a positive temperature gradient and extends up to the stratopause at about 45-55km. This is followed by the *mesosphere* where the temperature declines up to a height of 80-85km. Beyond this lies the *thermosphere* where the general temperature trend is an increase, depending heavily on solar activity for its precise variation. A fluid description is viable up to around 600km, by which point the atmosphere is becoming very diffuse and particles can achieve escape velocity.

Although we will pay most attention to the temperature inhomogeneity, none of the other physical parameters are constant throughout the entire atmosphere either. Incoming radiation from the Sun not only determines the temperature of the Earth's atmosphere but also contributes to the ionisation of the neutral particles, providing a second classification scheme. Since we will not be concerned with charged particle effects, we simply note that the lowest ionospheric region, the *D*-region extends upwards from around 60km. The mean molecular weight is constant up to about 80km but declines steadily above this height. The *turbopause* (~105km) signifies the boundary between the turbulent regions of the atmosphere and those where laminar flow always exists.

Superimposed on this temporally constant background are the oscillations in which we are interested - atmospheric waves. Visual evidence for these waves is afforded by mountain wave clouds (clouds formed when gravity waves force oscillations in the saturated air near mountain tops) whilst pressure and density variations allow the observations to be quantified. Tropospheric gravity waves are observed to have periods in the range 5-10 minutes. Fluctuations in the atmosphere above the troposphere were often ascribed to turbulent effects until the suggestion by Hines (1960) that these were, in fact, internal gravity waves. He was led to his conclusions by photographic and radar studies of the deformation, by rapidly varying winds, of long-enduring meteor trails. The predominant characteristics of these irregular winds and their explanation in terms of gravity waves are listed below.

1) The irregular wind components exhibit strong variations in vertical distances of a few kilometres, leading to estimates for the vertical wavelength of the wave to be around 12km.

2) The dominant winds persist over long intervals with little change over tens of minutes of time, corresponding to wave periods of the order of 200 minutes. This is consistent with gravity waves which exist for frequencies below a certain limiting frequency (the Brunt-Väisälä frequency discussed in the next section). They therefore have a lower limiting period, a typical value for which is

$$\tau_g = \frac{2\pi c}{g\,(\gamma-1)^{1/2}} \approx 5\text{mins},$$

which is safely below the 200 minutes threshold for the period of the observed waves.(c is the sound speed, g is the gravitational acceleration and γ is the ratio of the specific heats at constant volume and pressure respectively.)

The first two points are sufficient to define the parameters of the internal gravity wave and we must now verify that gravity waves with these characteristics could also be responsible for the remaining observations.

3) The horizontal scale size of these winds exceeds the dominant vertical scale size by a factor of 20 or more. Asymptotically, the dispersion relation may be expressed as $\lambda_x/\lambda_z \approx \tau/\tau_g$. Inserting the values for vertical wavelength and period inferred from points 1) and 2) above, we find that $\lambda_x \approx 490$km which is adequate to meet property 3).

4) The dominant motions are nearly horizontal. The relevant asymptotic relation in this case, between velocities and wavenumbers is, $u_x/u_z \approx \lambda_x/\lambda_z$ and from the preceding paragraph this ratio is found to be about 40. Theory therefore predicts that vertical motions would be negligible when compared with those horizontally and this is again in accord with the observational evidence.

5) The speed of the dominant irregular wind tends to increase with height as does the dominant scale size. The scale size of internal gravity waves is also found to be amplified with height and this feature agrees with the final observation.

2.2. The solar atmosphere

At the same time as the great upsurge of meteorological interest in acoustic-gravity waves, there arose a similar interest in the solar physics community, initiated by the discovery of the solar 5-minute oscillation by Leighton (1960) (cf. Leighton et al., 1962). Several mechanisms were suggested to generate these oscillations - all with the common theme of atmospheric waves. It has also been postulated that atmospheric waves supply the non-radiative energy needed to heat the chromosphere and corona.

As well as the influence of temperature in determining the allowed wave modes, the Sun's magnetic field plays a far greater role than the Earth's since it can be many times stronger. The Sun's magnetic field is concentrated in certain *active* regions, where the field strength may be up to 10^{-1}T providing energy for a range of dynamical processes, and which are surrounded by areas of quiet Sun which are almost devoid of magnetic field (~10^{-4}T). A division of the Sun's atmosphere into layers with similar temperature characteristics (where the temperature signature of each layer is simply a manifestation of the different physical processes occurring within it) must therefore be combined with an examination of its magnetic field structure (Durrant, 1988).

Underlying the atmosphere is the solar interior where many of the atmospheric effects originate. The outermost layer of the solar interior, the *convection zone* which ranges from 0.7 to 1 solar radius, is so-called because it is convectively unstable. It is the convective motions in this region which lead to the observed *granulation* of the neighbouring *photosphere*. The granules form an irregular cellular pattern of polygonal bright elements with horizontal scales of about 1400km, surrounded by a network of dark lanes. These intergranular lanes contain the magnetic elements which are swept to the edge of the granules by local motions. The solar surface is taken to occur at the height where the gas changes sharply from being opaque to transparent as radiative transport gains overwhelmingly in efficiency and the lowest of the atmospheric layers, the photosphere, is the first we can actually see. In the photosphere, the temperature is relatively low (averaging 6000°K) which causes the scaleheight to be small and results in a rapid exponential decrease of density with height. The relative importance of the magnetic field must thus increase with height since the balance of thermodynamic to magnetic pressure, $\beta = p/(B_0^2/2\mu_0)$, has decreased. The equivalence of these two pressures is achieved at a point in the *chromosphere* and above this height wave motions are studied under the MHD approximations and not those applicable to the Earth's neutral atmosphere. Near the photosphere-chromosphere boundary, around 500km above the solar surface, the temperature achieves its absolute minimum value (≃ 4300°K). In the chromosphere, the temperature increases, slowly at first and then in an irregular fashion to around 2.5×10^{4}°K. In the *transition* region there is an extremely rapid change of properties over a very short distance - a temperature rise to 10^{6}K in a few hundreds of kilometres. The temperature variation in the photosphere and chromosphere is illustrated in Figure (4.2).

Together, these three regions form a relatively thin layer, less than 3×10^{6}m thick, while the *corona*, and its extension the *solar wind*, continue out to the planets and beyond. The one-component fluid equations may be used below the transition region since the oscillation frequencies are much lower than the collisional frequency and so plasma effects are negligible. Viscosity can be neglected because the ratio of inertial to viscous acceleration (the Reynolds number) is $\approx10^{9}$. Thermal conductive flux is of the order of 10^{-9} times the mechanical flux in the chromosphere and below so that thermal conduction only becomes appreciable in the corona. As in the terrestrial case, we are usually interested in waves with sufficiently long periods to enable us to ignore rotational effects. When the effect of the magnetic field is included (magneto-acoustic-gravity waves), the electrical conductivity is high enough to be assumed perfect and hence the magnetic field lines are frozen in. Although radiative dissipation may modify the waves, it does not introduce any new wave types and is usually ignored. The equations of magnetoionic theory accurately model conditions almost everywhere in the solar corona and so are used to describe this region in preference to the fluid equations.

2.3. Implications for the model

From this brief summary of the properties of the terrestrial and solar atmospheres, it is obvious that no simple model can be constructed which is globally correct. Either we must resort to numerical simulations including a realistic profile for all the parameters or we must restrict ourselves to a more limited analytic approach. The best we can hope for is to construct a reasonable model of a particular area. For instance, by taking a fixed linear temperature gradient, we could satisfy conditions in the stratosphere *or* the mesosphere but not both. Similarly, results obtained for a neutral atmosphere may be applicable to conditions in the quiet photosphere but not in the upper chromosphere where magnetic effects are no longer negligible. Studies of the quiescent Sun can provide quite satisfactory models despite ignoring the magnetic field entirely. We will show in later sections that even the smallest deviation from homogeneity produces sufficient complication to render analytic solutions impossible in the majority of cases.

We will begin with a description of the wavemodes which can exist in an isothermal atmosphere, starting with the natural resonant frequency - the Brunt-Väisälä frequency mentioned above.

3. The Brunt-Väisälä frequency

An unmagnetised plasma, displaced slightly from its equilibrium position, will oscillate at its natural frequency, $\omega_{pe}^2 = n_0 e^2/\varepsilon_0 m_e$, as shown in Chapter 1. When the electrons move relative to the stationary ions, a potential difference is created which causes the electrons to accelerate back towards their initial positions. They tend to overshoot and are pulled back in the opposite direction, thus sustaining the oscillation. The fields in this case are self-induced and not externally imposed. An atmosphere under the force of gravity exhibits a similar resonant frequency called the *Brunt-Väisälä* frequency (Väisälä, 1925; Brunt, 1927). In the plasma, the energy of the system alternates between the kinetic energy of the moving electrons and the electric field potential while in the gravitationally stratified atmosphere, the atmospheric fluid elements transfer their energy to the gravitational potential. A significant difference between these two natural frequencies is that the plasma oscillations are longitudinal (parallel to the electric field) but the waves supported by gravity are transverse (with respect to the velocity field). A second difference is that electron plasma waves are a high frequency phenomenon (this is why the heavy ions with their large inertia may be omitted from the calculation) whereas gravity waves occur at relatively low frequencies (at higher frequencies, the gravitational restoring forces are negligible compared to those due to compressibility). In our analysis, we will treat both plasmas and atmospheres as compressible fluids subject to different restoring forces and employ the fluid equations of Chapter 3.

In order to derive the Brunt-Väisälä frequency explicitly, we examine the behaviour of a parcel of air in the atmosphere, displaced slightly from its equilibrium (which is assumed to be stable). Let the parcel be displaced at an angle θ to the vertical and the density difference between the parcel and its new surroundings be $\delta\rho$. Equating the forces acting on the air parcel after it has moved, the inertial force is balanced by a buoyancy force due to the pressure (density) difference between the parcel and its surroundings:

$$\rho \frac{\partial^2 z}{\partial t^2} = -\delta\rho \, g \cos\theta.$$

By the equation of state for a perfect gas, $p = \rho RT/M$, where $R = 8.31\times10^3 \mathrm{J} kmol^{-1}\mathrm{K}^{-1}$ is the universal gas constant (related to Boltzmann's constant and Avogadro's number by $R = k_B N_a$) and M is the mean molecular weight expressed in $kg/kmol$. Thus the temperature and density are related by $\delta\rho/\rho = \delta T/T$ if hydrostatic equilibrium is maintained. If there is no energy exchange between the parcel and its surroundings, the displacement is adiabatic and δT is given by the product of the difference between the adiabatic and atmospheric pressure gradients multiplied by the vertical component of the displacement.

Substituting this in the complete expression, we find:

$$\frac{\partial^2 z}{\partial t^2} = -\omega^2 z = \frac{1}{T}\left(\frac{dT}{dz}\bigg|_{\text{adiab}} - \frac{dT}{dz}\bigg|_{\text{atmos}}\right) g \cos^2\theta \, z ,$$

which represents waves satisfying the highly anisotropic dispersion relation, $\omega^2 = \omega_B^2 \cos^2\theta$. This defines the Brunt-Väisälä frequency to be $\omega_B^2 = \frac{g}{T}(dT/dz|_{\text{atmos}} - dT/dz|_{\text{adiab}})$.

We temporarily adopt the notation $\alpha = \partial T/\partial z$ for the temperature gradient. The adiabatic temperature gradient, α^*, may be used to gauge the stability of the atmosphere. Suppose we consider first an atmosphere whose temperature gradient exceeds the adiabatic temperature gradient. Then a parcel of air displaced adiabatically upwards from the position in the atmosphere where $\alpha = \alpha^*$ will follow the adiabatic temperature gradient. Hence, it will be at a lower temperature than its surroundings and will sink back to its starting point. An air packet displaced downwards will similarly rise and the situation is stable. If $\alpha < \alpha^*$, any displaced air will keep moving away from its original position and the situation is unstable. Now

$$\omega_B^2 = \frac{g}{T}(\alpha - \alpha^*),$$

so that for an atmosphere with $\alpha < \alpha^*$, the Brunt-Väisälä frequency becomes imaginary and the motions are no longer oscillatory but increase exponentially in time and the atmosphere is convectively unstable. Thus internal gravity waves and convection can be regarded as the stable and unstable responses to the same restoring force.

The Brunt-Väisälä frequency in an isothermal atmosphere is $\omega_g^2 = -g\alpha^*/T$ since the temperature gradient vanishes. From the equation of state for a perfect gas, we have

$$\frac{dT}{dz} = \frac{T}{p}\left(\frac{dp}{dz} - \frac{p}{\rho}\frac{d\rho}{dz}\right).$$

For an adiabatic expansion, $p\rho^{-\gamma} =$ constant, so that $dp/d\rho = \gamma p/\rho$ and the above expression gives

$$\alpha^* = \frac{T}{p}\left(1 - \frac{1}{\gamma}\right)\frac{dp}{dz} = \frac{T}{p}\frac{(1-\gamma)}{\gamma}\rho g ,$$

where we have used the fact that the displacement does not disturb the hydrostatic equilibrium. Thus

$$\omega_g^2 = \frac{(\gamma - 1) g^2}{c^2},$$

which is the form of the isothermal Brunt-Väisälä frequency we will use most often. In an atmosphere with a non-zero temperature gradient, the Brunt-Väisälä frequency is given by $\omega_B^2 = \omega_g^2 + \frac{g}{c^2}\frac{dc^2}{dz}$.

4. The model equations

Although this model at first sight appears to be greatly different from the cold plasma model of Chapters 5 and 6, there is a significant area of common ground between them. First, as we showed in Chapter 3, they share a common derivation - either from kinetic theory or from physical arguments involving the conservation of flux of quantities such as density. Thus the model equations are basically the same, with different approximations applied to each.

Secondly, the wave motion in each case is governed by external forces - here by gravity and in the following chapters by a magnetic field - which define *preferred directions* in their environment and so cause the plasma, in one case, and the atmosphere in the other to be highly anisotropic. The complications inherent in calculations involving magnetic fields are avoided in this chapter, thus considerably simplifying the geometry. It is therefore an ideal starting point for studies of waves in inhomogeneous media.

Our primary aim is to study the propagation of waves through a medium with non- constant parameters and, if possible, to construct a framework within which energy transfer between the eigenmodes of the system may be studied. Our first objective is therefore to identify the wave modes which describe the system.

As discussed qualitatively above, sound waves and gravity waves (the two extremes of frequency for acoustic-gravity waves) may exist in the atmosphere of a planet or star, but there is also the possibility of additional wavemodes. Because the Earth is not fixed in space but rotates about an axis with angular velocity Ω_E, it must also be subject to two additional forces due to its angular velocity. The centrifugal force, given by $\Omega_E \times (\Omega_E \times \mathbf{r})$ has a maximum value of $\Omega_E^2 R_E = 3.38 \times 10^{-2} ms^{-2}$ which is much smaller than the acceleration due to gravity and may therefore be ignored. The Coriolis force, $\mathbf{F}_{\mathrm{Cor}} = -2m_E \Omega_E \times \dot{\mathbf{r}}$ (where $\dot{}$ denotes $\partial/\partial t$), acts at right angles to the motion. Its importance lies not in its strength, because it is comparatively weak, but in the fact that it may produce a large deflection in motions with large time scales. Henceforth, we will use the equations of motion outlined in the following section and will neglect the Coriolis term. This is equivalent to avoiding long period waves (very low frequency) waves where the rotational effects are most significant and restricting our domain of interest to oscillations at higher frequencies.

We use the fluid equations to describe our model, taking the atmosphere to be a non-viscous, compressible fluid containing a gravitational field and not subject to thermal losses. These are the equations of mass, momentum and energy conservation which together form a system of 5 equations in 5 unknowns:

$$\frac{D\rho}{Dt} = -\rho\nabla\cdot\mathbf{u}, \tag{4.1}$$

$$\rho\frac{D\mathbf{u}}{Dt} = -\nabla p + \rho\mathbf{g}, \tag{4.2}$$

$$\frac{Dp}{Dt} = c^2\frac{D\rho}{Dt}, \tag{4.3}$$

where ρ, the mass density, $\mathbf{u}$, the fluid velocity and p, the fluid pressure are the unknowns. We denote the sound speed squared (a specified function of height, z through its dependence on the temperature) by c^2 and $\mathbf{g}$ represents the constant acceleration due to gravity.

To complete the description of the system, we must also specify our assumptions about the equilibrium. First, we take the equilibrium to be constant in time. The assumption of a static equilibrium means that we may set $\partial\rho_0/\partial t = \partial p_0/\partial t = 0$. Second, we assume that there is no wind in equilibrium, i.e. $\mathbf{u}_0 = 0$. A non-zero $\mathbf{u}_0$ produces a Doppler shift of the frequency and inhomogeneities in this shifted frequency have been used by some authors (Chimonas and Hines, 1986) to explain ducting of atmospheric waves.

From the equilibrium form of the momentum equation, we demonstrate how the natural stratification of the atmospheric pressure and density arises. Inserting the assumptions of a static equilibrium with $\mathbf{u}_0 = 0$ in the momentum equation, (4.2) yields the equation of hydrostatic equilibrium,

$$\nabla p_0 = \rho_0\mathbf{g},$$

or, in one dimension:

$$\frac{\partial p_0}{\partial z} = -\rho_0 g,$$

where the minus sign arises because the force of gravity increases downwards whilst altitude is measured upwards. This equation expresses the fact that the pressure gradient must be equal and opposite to the gravitational force in order to maintain a balanced state. We use the equation of state for a perfect gas to eliminate the density and integrate to obtain the pressure as a function of height:

$$p_0 = p_g \exp\left(-\int_0^z \frac{Mg}{RT}\,ds\right),$$

where p_g is the pressure at $z = 0$. It is useful to define a quantity, $H = RT/Mg = c^2/\gamma g$ which is the scaleheight of the neutral atmosphere, the height in which the pressure will fall to $1/e$ of its original value. For an isothermal atmosphere, H is constant and both pressure and density fall off exponentially with height:

$$\frac{p_0}{p_g} = \frac{\rho_0}{\rho_g} = \exp(-\frac{z}{H}) = \exp(-\frac{\gamma g z}{c^2}),$$

but in practice, none of M, g and T are constant, so that the scaleheight varies. Even the ratio of specific heats, γ, is not strictly constant. Since the acceleration due to gravity only varies by 5% over all latitudes and heights of up to two hundred kilometres and the mean molecular weight is identically constant for the first 80km with only a slow decrease thereafter, these two effects are neglected and any inhomogeneity in H will be attributed purely to a gradient in temperature.

We proceed in the usual way by linearising equations (4.1) to (4.3) and inserting the equilibrium conditions where appropriate to get :

$$\frac{\partial \rho_1}{\partial t} + \mathbf{u}_1 \cdot \nabla \rho_0 = -\rho_0 \nabla \cdot \mathbf{u}_1, \qquad (4.4)$$

$$\rho_0 \frac{\partial \mathbf{u}_1}{\partial t} = -\nabla p_1 + \rho_1 \mathbf{g}, \qquad (4.5)$$

$$\frac{\partial p_1}{\partial t} + \mathbf{u}_1 \cdot \nabla p_0 = c^2 (\frac{\partial \rho_1}{\partial t} + \mathbf{u}_1 \cdot \nabla \rho_0). \qquad (4.6)$$

At this stage, further simplification of the system may be performed by Fourier transforming these equations in time and space but, as discussed in Chapter 2, this is only useful when the coefficients of equations (4.4) to (4.6) are constant. Alternatively, we combine the three linearised equations into a single ode in one of the physical parameters (we will later choose u_z).

We first set the temperature gradient to be zero and perform the Fourier analysis in order to construct an algebraic dispersion relation relating the wavenumber and frequency of the allowed modes of the homogeneous system. This provides us with a comparison for the inhomogeneous case.

5. The dispersion relation

The dispersion relation for acoustic and gravity waves in an isothermal atmosphere has been derived by a number of authors (Eckart, 1960; Gossard and Hooke, 1975). Because of the stratification of the equilibrium pressure and density, even the most elementary treatment must tackle the problem of position-dependent coefficients. By appropriate manipulation, all height dependence may be incorporated in the variables, permitting full Fourier transformation. We follow the approach of Beer (1974) and rearrange equations (4.4) to (4.6) so that our set of dependent variables becomes $\mathbf{u}_1$, ρ_1/ρ_0 and p_1/p_0. Henceforth we will drop the subscript on the perturbed velocity since there is no equilibrium velocity with which to confuse it.

To eliminate the gradient of the equilibrium density from equation (4.4), we use the definition of the sound speed:

$$\nabla p_0 = \nabla(\frac{\rho_0 c^2}{\gamma}) = \frac{c^2}{\gamma}\nabla\rho_0 + \frac{\rho_0}{\gamma}\nabla c^2.$$

In the case of the homogeneous atmosphere which we are considering, the temperature is a constant and therefore the gradient of the sound speed is zero. Then the above expression becomes $\nabla\rho_0 = \gamma\nabla p_0/c^2$ and all the terms involving the derivatives of the equilibrium pressure can be expressed as functions of the acceleration due to gravity using the equation for hydrostatic equilibrium. We also wish to eliminate any gradients in the perturbed pressure in favour of a derivative of the ratio of first to zero order pressures, by using:

$$\nabla(\frac{p_1}{p_0}) = \frac{1}{p_0}\nabla p_1 - \frac{p_1}{p_0^2}\nabla p_0.$$

Thus, on rearranging equations (4.4) to (4.6) and using the substitutions outlined above, we obtain:

$$\frac{\partial(\rho_1/\rho_0)}{\partial t} + \nabla\cdot\mathbf{u} + \frac{\gamma}{c^2}\mathbf{u}\cdot\mathbf{g} = 0, \tag{4.7}$$

$$\frac{\partial\mathbf{u}}{\partial t} - \frac{\rho_1}{\rho_0}\mathbf{g} + \frac{c^2}{\gamma}\nabla(\frac{p_1}{p_0}) + \frac{p_1}{p_0}\mathbf{g} = 0, \tag{4.8}$$

$$\frac{\partial(p_1/p_0)}{\partial t} - \gamma\frac{\partial(\rho_1/\rho_0)}{\partial t} - \frac{\gamma(\gamma-1)}{c^2}\mathbf{u}\cdot\mathbf{g} = 0. \tag{4.9}$$

We now Fourier transform in space and time, so that all perturbed quantities have $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$ dependence, and we rotate the horizontal axes in order to define a coordinate system in which, for convenience, there is no y-dependence (i.e. $\partial/\partial y = 0$). The second component of the momentum equation, (4.8), then yields $i\omega u_y = 0$ and we see that the velocity does not have a $y-$ component. Since the coefficients of equations (4.7) to (4.9) are constant, the variation of the remaining non-zero variables is taken to be sinusoidal and their amplitudes and phases are related by:

$$\frac{u_x}{X} = \frac{u_z}{Z} = \frac{\rho_1/\rho_0}{R} = \frac{p_1/p_0}{P} = Ae^{i(K_x x + K_z z - \omega t)},$$

where ω is real, K_x and K_z are complex wavenumbers and X, Z, R and P are complex amplitudes. By selecting only real ω, we are limiting ourselves to only stable oscillations whilst permitting the wavenumbers to have imaginary parts allows both evanescent and oscillatory motion to be studied. Using these definitions, equations (4.7) to (4.9) may now be written in matrix notation as:

$$\begin{bmatrix} -i\,\omega & 0 & 0 & i\,K_x c^2/\gamma \\ 0 & -i\,\omega & g & i\,K_z c^2/\gamma - g \\ i\,K_x & i\,K_z - \gamma g/c^2 & -i\,\omega & 0 \\ 0 & \gamma(\gamma-1)g/c^2 & i\,\gamma\omega & -i\,\omega \end{bmatrix} \begin{bmatrix} X \\ Z \\ R \\ P \end{bmatrix} = 0 \qquad (4.10)$$

which describes a set of 4 homogeneous simultaneous equations in 4 unknowns. There therefore exists a unique non-trivial solution if and only if the determinant of the 4×4 matrix is zero, i.e. when:

$$\omega^4 - \omega^2 c^2 (K_x^2 + K_z^2) + (\gamma-1) g^2 K_x^2 - i\,\omega^2 \gamma g K_z = 0. \qquad (4.11)$$

This is the general dispersion relation for wave propagation in an isothermal atmosphere under the approximations and assumptions stated above. From equation (4.10), we may also derive relations between the amplitudes of each variable - often called the *polarisation relations.* From the first row of (4.10), we find that $P/X = \gamma\omega/K_x c^2$ which may be substituted into the second row to eliminate P. We now solve the remaining pair of simultaneous equations as if Z were known to obtain the relative ratios of R and Z. In his classic paper, Hines (1960) chose the denominator of the resulting expression to define $Z = \omega^3 - \omega K_x^2 c^2$. This choice fixes the remaining amplitudes to be:

$$P = \gamma\omega^2 K_z + i\frac{\gamma g\,\omega^2}{c^2},$$

$$R = \omega^2 K_z + i\frac{\gamma g\,\omega^2}{c^2} - i\,(\gamma-1) g K_x^2,$$

$$X = \omega K_x K_z c^2 + i\,\omega g K_x.$$

There are several conclusions which we may draw immediately from equation (4.11). First of all we show that the general wavenumbers K_x and K_z may not both be real and non-zero.

Proof

Suppose that $K_x = \alpha + i\beta$ and $K_z = \psi + i\delta$ (where α, β, ψ, $\delta \in \mathbb{R}$). We divide equation (4.11) into its real and imaginary parts:

Real part

$$\omega^4 - \omega^2 c^2(\alpha^2 - \beta^2 + \psi^2 - \delta^2) + (\gamma - 1)g^2(\alpha^2 - \beta^2) + \omega^2\gamma g\,\delta = 0.$$

Imaginary part

$$-2\omega^2 c^2(\alpha\beta + \psi\delta) + 2(\gamma - 1)g^2\alpha\beta - \omega^2\gamma g\,\psi = 0.$$

Suppose that both K_x and K_z are real, i.e. $\beta = \delta = 0$, then the imaginary part of equation (4.11) becomes

$$\omega^2\gamma g\,\psi = 0,$$

forcing $\psi = 0$ and thus the vertical wavenumber, K_z, must also be zero. As was asserted above, imposing a real horizontal wavenumber forces any propagation in the vertical direction to have a growing or decaying envelope.

We may also use the division of equation (4.11) into real and imaginary parts to remove the undesirable i term from the dispersion relation. For a real horizontal wavenumber ($K_x = k_x$), we know that the vertical wavenumber must be of the form $K_z = k_z + i\kappa$ where k_z and κ are also real. Now the real part of equation (4.11) becomes:

$$\omega^4 - \omega^2 c^2(k_x^2 + k_z^2 - \kappa^2) + (\gamma - 1)g^2 k_x^2 + \omega^2\gamma g\,\kappa = 0,$$

and its imaginary part becomes:

$$-2\omega^2 c^2 k_z\kappa - \omega^2\gamma g k_z = 0.$$

We see from the second of these relations that either the real part of the vertical wavenumber is zero or we obtain an equation for the vertical growth factor:

$$\kappa = -\frac{\gamma g}{2c^2}.$$

Generally, exponential growth of a wave over more than a finite distance would contravene energy conservation but in the case of the atmosphere, this constant growth is permitted by the ambient density stratification. If we consider the total wave energy density (neglecting potential energy), we have

$$\varepsilon = \tfrac{1}{2}\rho_0 u^2,$$

and we have already shown that $\rho = \rho_g \exp(-\gamma g z / c^2)$ so that energy conservation demands the form of the imaginary part of the vertical wavenumber to be the same as that derived from the dispersion relation above.

We substitute the above form of κ into the real part of the dispersion relation to construct a final version which contains no imaginary terms and which will yield the real part of the vertical wavenumber in terms of the frequency and horizontal wavenumber:

$$\omega^4 - \omega^2 c^2 (k_x^2 + k_z^2 + (\frac{\gamma g}{2c^2})^2) + (\gamma - 1) g^2 k_x^2 = 0. \quad (4.12)$$

Now we examine the regions in which we may find real solutions for the dispersion relation. Considering (4.12) to be quadratic in ω^2, we write:

$$2\omega^2 = c^2 (k_x^2 + k_z^2 + (\frac{\gamma g}{2c^2})^2) \pm (c^4 (k_x^2 + k_z^2 + (\frac{\gamma g}{2c^2})^2)^2 - 4(\gamma - 1) g^2 k_x^2)^{½} \quad (4.13)$$

To examine the zero-wavelength limit, we take an asymptotic expansion of the solutions of equation (4.12) and so rewrite equation (4.13) in a suitable form:

$$2\omega^2 = c^2 (k_x^2 + k_z^2) + (\frac{\gamma g}{2c})^2$$

$$\pm c^2 k_x^2 (1 + \frac{2k_z^2}{k_x^2} + \frac{k_z^4}{k_x^4} + \frac{\gamma^2 g^2}{2k_x^2 c^4}(1 + \frac{k_z^2}{k_x^2}) + (\frac{\gamma g}{2k_x c^2})^4 - \frac{4(\gamma - 1) g^2}{k_x^2 c^4})^{½} .$$

For simplicity, we consider $k_z = 0$, and see that as $k_x \to \infty$, the larger root takes the form of a sound wave, $\omega^2 = c^2 k_x^2$. Henceforth we will therefore identify this branch as the *acoustic mode*. The other root approaches a wave with constant frequency $\omega_g^2 = (\gamma - 1) g^2 / c^2$ which is the isothermal Brunt-Väisälä frequency, the resonant frequency of the atmosphere which was derived from first principles in an earlier section and we thus designate this root a *gravity wave*.

The points where the wavelength becomes infinite, i.e. $k_x \to 0$, may be studied in a similar fashion but we use a more appropriate form for the solutions, by writing equation (4.13) as:

$$2\omega^2 = c^2 (k_x^2 + k_z^2) + (\frac{\gamma g}{2c})^2 \pm (\frac{\gamma g}{2c})^2$$

$$(1 + (\frac{2c^2}{\gamma g})^4 (k_x^2 + k_z^2)^2 + 2(\frac{2c^2}{\gamma g})^2 (k_x^2 + k_z^2) - \frac{64(\gamma - 1) k_x^2 c^4}{\gamma^4 g^2}))^{½} .$$

We set $k_z = 0$ and examine the behaviour of equation (4.12) as $k_x \to 0$:

$$\omega^2 \to (\frac{\gamma g}{2c})^2 \pm (\frac{\gamma g}{2c})^2 (1 + O(k_x^2)).$$

As indicated in Chapter 1, a *cutoff* occurs in a medium supporting wave motion when the wavelength becomes infinite (the wavenumber goes to zero) and, conversely, a *resonance* occurs when the wavelength becomes zero, so that the cut-offs of the lower and upper branches of the solution occur at $\omega = 0$ and $\omega_a = \gamma g/2c$ (the *acoustic cutoff frequency*) respectively.

For $k_z \neq 0$, the general behaviour of the modes is unaltered but the values of the acoustic cutoff frequency and the zero-wavelength limit of the gravity mode are increased and decreased respectively with the result that the region where propagating solutions are forbidden increases in size. Any of the properties of equation (4.12) can best be examined graphically. In Figure (4.3), we plot frequency against horizontal wavenumber for four different values of k_z, including $k_z = 0$. The upper curve represents acoustic waves and demonstrates the acoustic cutoff frequency for $k_x \rightarrow 0$ and asymptotically approaches a straight line, $\omega = ck_x$. The lower curve vanishes at the origin and in the short wavelength limit tends to the Brunt-Väisälä frequency. Between these two curves lies the evanescent region where no waves are possible. This gap between ω_a and ω_g for which no frequency may propagate with a purely real horizontal wavenumber is clearly visible. As k_z increases, the two curves are seen to move apart, so that the evanescent region becomes larger. Figure (4.4) depicts ω versus k_z for a set of values of k_x. (Note that for $k_x = 0$ the gravity wave branch would disappear and only acoustic waves could propagate.)

Only one evanescent wave will prove to be of interest to us later - the Lamb wave which satisfies the dispersion relation $\omega^2 = c^2 k_x^2$ and, of course, $k_z = 0$. Substituting these values in the real part of equation (4.11), we find that the imaginary part of the vertical wavenumber must satisfy:

$$\kappa = \frac{g}{2c^2}(-\gamma \pm (\gamma - 2)).$$

Lamb waves are always evanescent, lying in the non-propagating region of Figure (4.3). Unlike other evanescent waves, they are linearly, not elliptically, polarised. We demonstrate this by substituting their dispersion relation in the polarisation relations to find that $Z = \omega^3 - \omega c^2 k_x^2 = 0$. Thus these waves do not have any vertical velocity variation and are linearly polarised. The remaining polarisation relations of both Lamb waves are therefore found from:

$$\begin{bmatrix} -i\omega & 0 & i\omega c/\gamma \\ i\omega/c & -i\omega & 0 \\ 0 & i\gamma\omega & -i\omega \end{bmatrix} \begin{bmatrix} X \\ R \\ P \end{bmatrix} = 0,$$

so that the density, pressure and x-component of the velocity have amplitudes related by

$$R = 1 \qquad P = \gamma \qquad X = c.$$

6. Non-uniform atmosphere

We have so far only been concerned with the analysis of a uniform atmosphere. Although the above analysis is not applicable to the case of an atmosphere with a temperature gradient (or any such inhomogeneity), it serves as a comparison for what follows and is useful for providing physical insight into the non-uniform case. As soon as inhomogeneities are introduced, a host of new effects arise - cf. Francis (1975) for a review of atmospheric waves in uniform and non-uniform media. The effect which has received most attention in the past has been the possibility of trapping waves. Non-uniform properties may introduce new cutoff frequencies and hence additional reflection points in the medium. If waves are continuously reflected between two such reflection layers, a standing wave pattern will be set up where the waves in opposite directions interfere constructively. Thus a waveguide can be created which only supports a certain discrete set of eigenmodes. Models of ducting by temperature structure have been studied extensively by both the atmospheric and solar physics communities although there has been little interaction between the two. A second new phenomenon caused by spatial variation of the atmospheric parameters is that of mode conversion (introduced in Chapter 2), a subject which has received only cursory attention in the literature. We will therefore look for a situation where a propagating wave is incident on a region where it becomes evanescent, beyond which the wave can propagate once more. Under such conditions, some of the wave may not be transmitted but will instead be reflected.

On purely physical grounds, it is tempting simply to extend the results of the isothermal model by making the acoustic cutoff (ω_a) and gravity-wave resonance (ω_g) frequencies position-dependent (see the review of solar oscillations in Jordan (1981)). It could then be argued that gravity waves, with all but the smallest horizontal wavenumbers, would propagate only for frequencies below the non-isothermal Brunt-Väisälä frequency (the generalisation of ω_g^2):

$$\omega_B^2 = \omega_g^2 + \frac{g}{c^2}\frac{dc^2}{dz}.$$

We will demonstrate in a later section a more rigorous method of deriving the reflection levels for an atmosphere with a temperature gradient from a WKBJ approach. At that point, we will show that the above expression gives a reasonable estimate of the height at which reflection is likely to occur for gravity waves with large horizontal wavenumber.

There is no single correct ode describing any physical system and the choice of which one to use may depend on tractability or on the ultimate aim of the analysis. Two common choices in the study of waves in stratified atmospheres are to take either i) $\nabla \cdot \mathbf{u}$ or ii) $\rho_0^{1/2} \nabla \cdot \mathbf{u}$ as the dependent variable. These have the following advantages over other choices :-

1) The coefficients of the resulting differential equation do not contain height derivatives of the scaleheight (or the background wind), so that the wave variables are continuous across boundaries where the temperature or wind velocity change suddenly. (This point is particularly useful in numerical calculations.)

2) An extra singularity (at $\omega^2 = c^2 K_x^2$) is removed which simplifies solution.

3) The choice of variable ii) contains a factor to compensate for the growth in amplitude of the wave, which results from the exponential decrease in density with increasing height (cf. §5).

After examining some results obtained using variable choices i) and ii), we will proceed to investigate the "sonic line" singularity by deriving local solutions appropriate to the neighbourhood of this point.

Pitteway and Hines (1965) considered the possibility of inhomogeneities due to changes in height of either the temperature or the velocity of the (horizontal) background wind, $\mathbf{u}_0$, and showed that the linearised system of equations (4.4) to (4.6) could be expressed as a pair of coupled first order ode's in the variables $\xi = \nabla \cdot \mathbf{u}/w$ and $\zeta = K_x u_z/w$:

$$\xi' = \left(\frac{\gamma}{\alpha} - \frac{1}{\beta}\right)\xi - \left(\beta - \frac{1}{\beta}\right)\frac{\zeta}{\alpha} \qquad \zeta' = \left(1 - \frac{\alpha}{\beta}\right)\xi + \frac{\zeta}{\beta} \tag{4.14}$$

where $' = d/ds$, $s = K_x z$, $\alpha = \gamma H K_x$, $\beta = w^2/gK_x$, $H = c^2/\gamma g$ and $w = \omega - K_x u_{0x}$ is the Doppler-shifted frequency due to the horizontal background wind. In the case of constant (or absent) $\mathbf{u}_0$, β is constant and ζ can be eliminated from equations (4.14), leaving a second order ode:

$$\xi'' + \xi'\left(\frac{\alpha'}{\alpha} - \frac{\gamma}{\alpha}\right) + \xi\left(\frac{\alpha'}{\alpha\beta} + \frac{1}{\alpha}\left(\beta - \alpha + \frac{\gamma - 1}{\beta}\right)\right) = 0.$$

Transforming to the new variable $\phi = w\xi \exp\left(-\int \frac{\gamma}{2\alpha} ds\right)$ gives

$$\ddot{\phi} + \frac{\dot{H}}{H}\dot{\phi} + q^2\phi = 0, \tag{4.15}$$

where $\dot{} = d/dz$, we specify a real horizontal wavenumber, k_x, and

$$q^2 = k_x^2\left((\gamma - 1 + \gamma\dot{H})\frac{g^2}{\omega^2 c^2} - 1\right) + \frac{1}{c^2}\left(\omega^2 - \left(\frac{\gamma g}{2c}\right)^2\right).$$

Equation (4.15) was first derived by Martyn (1950). It is exact - terms involving $\dot{H}^2$ and $\ddot{H}$ vanish identically. An analytic solution exists only when the temperature (and hence c^2) has a very simple height dependence. Pitteway and Hines studied cases where temperature was an exponential ($\alpha = \alpha_0 \exp(As)$) or linear function of height ($\alpha = As$). They observed that the second order equation in ζ which is equivalent to equation (4.15) contains a singularity where $\alpha = \beta$ i.e. where $\omega^2 = c^2 K_x^2$ and so the equation in ζ is not as "convenient" to use. For the temperature profiles assumed, equation (4.15) becomes a form of Whittaker's confluent hypergeometric equation with solutions possible in terms of Laguerre functions. The authors then demonstrate that modes trapped by the entire atmosphere are possible if they satisfy the constraint that the wave amplitude tends to zero at the upper and lower bounds. From a series expansion, convergence arguments lead to the conclusion that only a discrete set of modes - indexed by n - can exist. Defining n as the order of the mode, an algebraic formula may be generated which replaces the simple dispersion as a description of the propagation characteristics.

A similar analysis was applied by Daniels (1967) to the thermosphere, using the field variable $\nabla \cdot \mathbf{u}$ in conjunction with an exponential temperature variation, defined by $c^2 = A + Be^{sz}$. The second order equation which he derived using the dimensionless independent variable $y = c^2/A$ was the Papperitz equation with solutions $y^{\lambda}(y-1)^{\mu}F(a,b,c;y)$ where λ and μ depend on the choice of frequency, horizontal wavenumber and temperature profile while $F(a,b,c;y)$ defines a hypergeometric function. Using energy considerations to fix the boundary conditions (again at $z = 0,\infty$), the complete set of eigenvalues was extracted. As well as perfectly guided modes, Daniels also noted the possibility of *leaky* modes with complex wavenumbers which resemble the freely propagating evanescent waves in an isothermal atmosphere.

The generation and propagation of waves in the Sun's atmosphere was considered by Moore and Spiegel (1964) who derived the isothermal dispersion relation, equation (4.12), and examined the limiting cases of vertical and nearly horizontal propagation in an atmosphere with varying temperature, under the WKBJ approximation. They found that the concept of acoustic cutoff frequency was generalised, via its dependence on sound speed in the case of purely vertical propagation. For nearly horizontal propagation, they discovered that the high frequency limit of gravity waves occurred at the non-isothermal Brunt-Väisälä frequency, ω_B. These results were extended to arbitrary directions of propagation, employing the generalised expressions to define the limits of wave propagation when the sound speed squared has a parabolic dependence on height (simulating conditions near the temperature minimum). As we will demonstrate later, this method does in fact result in a reasonable estimate of the height at which reflection of atmospheric waves can be expected in an inhomogeneous atmosphere.

In Chapter 2, we used the concept of quantum mechanical scattering from a potential in the context of wave propagation in an inhomogeneous medium. The approach hinged on expressing the appropriate second order ode in normal form, $y''+f(z)y = 0$, where $f(z)$ is the potential. The eigenvalues of the system, which define the relevant modes, were derived by considering the basic equations as a system of coupled first order ode's. A slightly different approach was adopted by Yu et al. (1980) and Tuan and Tadic (1982) who expressed their second order equations as a Sturm-Liouville equation:

$$\left(\frac{d}{dz}Q(z)\frac{d}{dz} - V(z) + \lambda^2\right)\psi = 0.$$

where $Q(z), V(z)$ are functions of ω, k_x and c^2, $\psi = \omega p_1/\rho_0$ and the eigenvalues are identified by $\lambda = k_x/\omega$. Again, the zeros of the potential represent the positions where reflection takes place. This procedure is handicapped by the labelling of eigenvalues by λ which is a constant and therefore includes a whole range of modes with different vertical dependence. The method outlined in Chapter 2 seems preferable since the potential is defined in what seems a more natural way and the eigenvalues, which vary with height, isolate the individual modes. Yu et al. generated solutions under both isothermal and slowly-varying (WKBJ) conditions, but the emphasis of their paper is on heuristic fitting of the eigenvalues to the scale of the potential well.

The consensus from all these investigations is that simple models may be used to demonstrate the likelihood of modes being ducted by atmospheric structure but that the simple models treated do not represent physically realistic systems. These models allow for two reflection points (at ground level and infinity) whereas in reality there may be several such levels throughout the atmosphere. The frustrating fact is that any model which is simple enough to have solutions in terms of standard functions does not admit the phenomena of interest. This conclusion has been deduced from studies examining ducting of modes but is equally true of the sparse mode coupling literature. Most discussions of linear wave coupling originate from numerical simulations.

Jones (1970) was the first to suggest that mode coupling occurs in the Earth's atmosphere during his examination of ionospheric oscillations. In this work, the author uses the terminology of electronics theory to examine a pair of coupled equations which are similar to the equations in ξ and ζ used by Pitteway and Hines (1965). The author invokes an atmospheric response function - the ratio of output (say u_z at some height in the atmosphere) to an input (some method of forcing at some other height) - which is dependent on frequency and horizontal wavenumber.

The results of Fourier transforming in space (x) and time may be multiplied by the response function to obtain the transformed output and an inverse Fourier transform taken to find the final spatial and temporal behaviour. Perturbations with frequency and wave number close to the normal mode produce an almost infinite response. As this is a numerical study, a full vertical thermal structure can be included and the author considers four different cases, each with a different background wind profile, $u_{0x}(z)$. The modal plots which result show modes coming together and then parting again, although the vertical wavenumber or its equivalent which defines individual modes is not stated explicitly. Jones terms any points where the approach is sudden *kissing modes* and those where it is more gradual *embracing modes*. These points are identified as places where modes transfer from one duct to another and therefore extend through large vertical distances.

In the context of ducted modes in the solar atmosphere, Christensen-Dalsgaard (1980) generated numerically similar mode diagrams demonstrating *avoided crossings* where the characters of two modes approach each other closely and then separate again. He interprets such events as a simple swapping of the properties of one type of wave to another at a point where the energy stored in the two wave types is comparable. He also suggests that an incoming wave of one species may be able to excite the second mode at such a point. As an expression of the concept of mode-conversion, this intuitive statement echoes the work of Cairns and Lashmore-Davies discussed in Chapter 2 where proximity of curves on a $k_z - z$ diagram indicated mode-conversion regions. It therefore has concomitant weaknesses but may be an indication that mode coupling is present.

7. The vector differential equation

Before we can resolve the problem of the extra singularity raised in the previous section, we must first construct a second order ode which contains this singularity from the model equations of §4. Although we will be interested in the effects of a temperature gradient, we include for completeness a short section on work which includes inhomogeneities caused by the presence of a magnetic field (and which consequently is aimed more at astrophysical applications). This section also provides an excellent illustration of the weaknesses inherent in employing local dispersion relations to describe non-uniform systems.

In order to generate a second order ode of the desired form, with the vertical velocity as dependent variable, we differentiate equation (4.5) to get:

$$\frac{\partial^2 \mathbf{u}_1}{\partial t^2} = -\frac{1}{\rho_0}\frac{\partial}{\partial t}(\nabla p_1) + \frac{\mathbf{g}}{\rho_0}\frac{\partial \rho_1}{\partial t},$$

and substituting for the perturbed pressure from equation (4.6) and for the perturbed density from equation (4.4):

$$= -\frac{1}{\rho_0}\nabla(-\rho_0(\mathbf{u}_1\cdot\mathbf{g}+c^2\nabla\cdot\mathbf{u}_1))+\frac{\mathbf{g}}{\rho_0}(-\mathbf{u}_1\cdot\nabla\rho_0-\rho_0\nabla\cdot\mathbf{u}_1)$$

$$= \nabla(\mathbf{u}_1\cdot\mathbf{g})+c^2\nabla\nabla\cdot\mathbf{u}_1+\nabla\cdot\mathbf{u}_1\nabla c^2+\frac{\nabla\rho_0}{\rho_0}(\mathbf{u}_1\cdot\mathbf{g}+c^2\nabla\cdot\mathbf{u}_1)-\mathbf{g}\nabla\cdot\mathbf{u}_1-\mathbf{u}_1\cdot\mathbf{g}\frac{\nabla\rho_0}{\rho_0},$$

where we have expanded out the ∇ terms and have used the identity $(\mathbf{a}\cdot\mathbf{b})\mathbf{c} = \mathbf{b}(\mathbf{a}\cdot\mathbf{c})+(\mathbf{b}\times\mathbf{c})\times\mathbf{a}$ together with the fact that the gravity vector is parallel to the density gradient. The definition of the sound speed in terms of the pressure and density plus the equation of hydrostatic equilibrium allow us to write $\nabla c^2 = \gamma\mathbf{g}-c^2\frac{\nabla\rho_0}{\rho_0}$. Thus the complete vector equation becomes (using the abbreviation $\dot{} = d/dt$):

$$\ddot{\mathbf{u}} = \nabla(\mathbf{u}\cdot\mathbf{g})+(\gamma-1)\mathbf{g}\nabla\cdot\mathbf{u}+c^2\nabla(\nabla\cdot\mathbf{u}). \qquad (4.16)$$

If the effects of gravity were negligible, equation (4.16) would reduce to the familiar equation describing the propagation of sound waves with velocity c, $\ddot{\mathbf{u}} = c^2\nabla(\nabla\cdot\mathbf{u})$.

8. Addition of a magnetic field

At this stage we digress to consider the effect on equation (4.16) of the introduction of a magnetic field. This addition permits the propagation of new modes for which the magnetic field provides the restoring force. The combined action of all three forces results in magneto-acoustic-gravity (MAG) waves which have received a great deal of attention from the solar physics community (Stein and Leibacher, 1974). Because of the greatly increased complexity of the system, analysis has concentrated on particular orientations of a magnetic field of constant magnitude and on specific angles of propagation. With a varying magnetic field, there is the additional requirement that $\nabla\cdot\mathbf{B}_0$ must be identically zero everywhere. In order to avoid partial differential equations, the field must therefore vary in a direction perpendicular to the plane in which it lies. Combining this with the restrictions already imposed by the conservation equations introduces further complication. An example of this inter-connection of equilibrium equations and the consequent limitations placed on a physical model will be found in Chapter 6. There, we will examine the effect of a spatially rotating magnetic field on wave propagation in the cold plasma model.

Before looking at the combined influence of magnetic and gravitational fields on a compressible magneto-fluid, we consider a plasma under the sole influence of a magnetic field. The frequency domain of interest makes the MHD regime most appropriate for use here. In a magnetic field the stresses are equivalent to a magnetic tension, B^2/μ_0, along the field lines and across, a hydrostatic pressure, $B^2/2\mu_0$. Since the latter can always be superimposed on the fluid pressure, the magnetic field lines behave effectively as elastic cords under a tension B^2/μ_0. In a perfectly conducting plasma, the particles behave as if they were tied to the magnetic field lines so that the lines of force act like mass-loaded strings under tension. By analogy with the transverse vibrations of elastic strings, we would therefore expect that, when the plasma is displaced slightly from equilibrium, the magnetic field lines would perform transverse vibrations at the velocity:

$$c_a = \left[\frac{\text{tension}}{\text{density}}\right]^{1/2} = \left[\frac{B^2}{\mu_0\rho}\right]^{1/2},$$

termed the *Alfvén velocity*, after Alfvén (1942). This is indeed the velocity of propagation of a transverse wave which propagates parallel to the magnetic field. This mode is known variously as the *shear Alfvén wave* or the *slow Alfvén wave*. For perpendicular propagation, the combined pressures cause a longitudinal wave, $\omega^2 = k^2(c_s^2 + c_a^2)$, the *compressional* or *fast* Alfvén wave, where we have used the notation c_s for the sound speed to avoid confusion.

With the introduction of a magnetic field into the acoustic-gravity wave problem, we require an additional evolution equation. This is the equation of conservation of magnetic flux, which may be written in linearised form as:

$$\frac{\partial \mathbf{B}}{\partial t} = (\mathbf{B}\cdot\nabla)\mathbf{u} - \mathbf{B}\nabla\cdot\mathbf{u}, \tag{3.13}$$

and is derived from Ohm's Law applied to a perfect conductor and the result then substituted into Faraday's Law. This statement also contains the assumption that there is no equilibrium velocity and that the ambient magnetic field is constant. The right hand side of the momentum equation, (4.2), also gains a term of the form $\frac{1}{\mu_0}(\nabla\times\mathbf{B})\times\mathbf{B}$ in the presence of a magnetic field. Combining the resulting set of equations produces:

$$\ddot{\mathbf{u}} = \nabla(\mathbf{u}\cdot\mathbf{g}) + (\gamma-1)\mathbf{g}\nabla\cdot\mathbf{u} + (c_a^2+c_s^2)\,\nabla(\,\nabla\cdot\mathbf{u}\,)$$
$$+\frac{1}{\mu_0\rho_0}((\mathbf{B}_0\cdot\nabla)^2\mathbf{u} - \mathbf{B}_0(\mathbf{B}_0\cdot\nabla)\nabla\cdot\mathbf{u} - (\mathbf{B}_0\cdot\nabla)\nabla(\mathbf{B}_0\cdot\mathbf{u})). \tag{4.17}$$

Even for an isothermal atmosphere and the constant magnetic field which we have assumed, equation (4.17) does not have constant coefficients. The density is stratified and this appears explicitly multiplying the final term and is implicit in the Alfvén velocity. In order to proceed, some authors derive the local dispersion relation (treating the Alfvén speed as constant) and observe that this divides into two factors. One represents an Alfvén wave whilst the other is a complicated fourth-order equation for k_z depending on ω, k_x, the magnetic field strength and orientation (McLellan and Winterberg, 1968). Special cases of magnetic field orientation and direction of propagation have been analyzed by different authors, including the case of vertical propagation (Bel and Mein, 1971).

The general case of an oblique magnetic field (lying in the x,z-plane) is treated in a series of papers by Zhugzhda and Dzhalilov (1984). These authors find that the complete system for arbitrary propagation in an oblique field, derived from equation (4.17) comprises a linear sixth-order ode with variable coefficients. Solution of the most general case with $k_y \neq 0$ is found to be impossible in terms of known functions. However, when $k_y = 0$, the Alfvén wave equation decouples from the system and the solution of the remainder is expressible in terms of Meijer and hypergeometric functions (Bateman Manuscript Project Staff, 1953). Because the magnetic field introduces a second preferred direction, there is no longer the degree of freedom available which previously allowed this simplification (rotation of the x,y-axes) without loss of generality. In the limit of a very weak or very strong field, the waves decompose into non-interacting modes whereas there is interaction and transformation of the waves in the region where $c_{az}^2 \approx c_s^2$. In the weak field region, the solutions tend to those for acoustic-gravity waves in an isothermal atmosphere with no field plus the slow magneto-acoustic mode in a weak field. For a strong field, the waves are those for slow magneto-acoustic modes in a strong field and depend on the relative signs of k_x and B_{0x}.

In the second paper of the series, the solutions derived in the first are used to examine the propagation and transformation of the waves (mode-conversion, in the language of Chapter 2) as they pass from a *weak field* region ($c_{az}^2 \ll c_s^2$) to a *strong field* region ($c_{az}^2 \gg c_s^2$). This problem is crucial to the development of the theory of solar atmospheric heating. By matching the coefficients of the asymptotic solutions at the boundary between the weak and strong field regions, and applying suitable energy constraints at infinity, it is possible to obtain reflection and transmission coefficients for every possible incident wave mode. The results indicate that the transformation of MAG waves depends strongly on the field inclination. It is also found that there are neither resonance levels nor valve effects. This contrasts with the work of Adam (1977) who applied local dispersion relation techniques to MAG waves in an oblique field and showed that two resonance levels arose, pertaining to waves running along or against the field inclination. Because each of the two levels only transmits in one direction, this has been described as a valve effect. This is one of the few examples where a direct comparison may be made between the results of a local dispersion relation applied to an inhomogeneous system and those of the complete analysis. The conclusion is the same as that of other authors who have performed similar comparisons (Diver, 1986). The disagreement between the two results can only be explained in terms of an inadequacy on the part of the approximate treatment. The dispersion relation does not reproduce the correct results and, in this case has added spurious effects, namely additional reflection levels, which are not observed by the more rigorous analysis.

9. The differential equation in u_z

Henceforth we will consider exclusively the case of inhomogeneity in the sound speed terms due entirely to temperature changing with height and will assume that the remaining parameters are constant while ignoring the effect of the magnetic field. We will therefore primarily be interested in the Earth's atmosphere, but the analysis is equally applicable to the Sun's atmosphere whenever the magnetic field is relatively unimportant (β is large). Since the coefficients have no time or horizontal space dependence, we may continue to assume sinusoidal variation with respect to these variables and may Fourier transform equation (4.16) in t and x. We assume that the horizontal wavenumber, k_x, is purely real throughout. As in the isothermal case, we take $\partial/\partial y = 0$ without loss of generality. Taking components of the vector equation (4.16) then yields:

$$-\omega^2 u_x = -ik_x g u_z - k_x^2 c^2 u_x + ic^2 k_x \frac{\partial u_z}{\partial z},$$

$$-\omega^2 u_y = 0,$$

$$-\omega^2 u_z = -\gamma g \frac{\partial u_z}{\partial z} - ik_x(\gamma - 1) g u_x + ic^2 k_x \frac{\partial u_x}{\partial z} + c^2 \frac{\partial^2 u_z}{\partial z}.$$

Notice that the choice of orientation of the horizontal wavenumber has resulted in the wind velocity having no y-component. We eliminate

$$u_x = \frac{ik_x(c^2 \partial u_z/\partial z - g u_z)}{(-\omega^2 + c^2 k_x^2)}$$

between the two remaining equations to obtain an equation in u_z only:

$$u_z'' S + u_z'(k_x^2 c^{2\prime} - \frac{\gamma g}{c^2} S) + u_z(\frac{S^2}{c^2} + \frac{(\gamma-1) g^2 k_x^2 S}{\omega^2 c^2} - \frac{k_x^4 c^{2\prime} g}{\omega^2}) = 0 \qquad (4.18)$$

where $'$ denotes $\partial/\partial z$ and we have defined $S = \omega^2 - c^2 k_x^2$. We see immediately that there is a singularity in this equation for any values of z where $\omega^2 = c^2 k_x^2$. In general, equation (4.18) has no solution in terms of standard functions, except in the cases where the temperature has a particularly simple height dependence (Pitteway and Hines, 1965).

Uchida (1965) also constructed the differential equation (4.18) while investigating the solar 5-minute oscillation. He attributed the observed oscillations to standing, compressional gravity waves trapped in a potential well caused by the temperature minimum in the low chromosphere. His numerical calculations showed that modes trapped by the assumed temperature structure would exhibit values of frequency and horizontal wavenumber close to those observed but current investigations favour a model where the cavity lies below the photosphere, since such a model succeeds in explaining more recent observations.

To express equation (4.18) in normal form, we use the transformation demonstrated for the general second order ode in Chapter 2. Here we set

$$u_z = \phi \exp(\tfrac{1}{2}\int(\frac{\gamma g}{c^2} - \frac{k_x^2}{S}\frac{dc^2}{dz})dz).$$

This gives $\dfrac{d^2\phi}{dz^2} + f(\omega, k_x; z)\phi = 0$ where the *potential* is of the form:

$$f(\omega, k_x; z) = \frac{S}{c^2} + \frac{(\gamma-1)g^2 k_x^2}{\omega^2 c^2} - \frac{\gamma^2 g^2}{4c^4} - (\frac{g k_x^4}{\omega^2 S} - \frac{\gamma g}{2c^2 S}(k_x^2 - \frac{S}{c^2}))\frac{dc^2}{dz}$$
$$- \frac{3k_x^4}{4S^2}(\frac{dc^2}{dz})^2 - \frac{k_x^2}{2S}\frac{d^2c^2}{dz^2}. \qquad (4.19)$$

This is the potential function calculated by Uchida which he plots for a selection of possible temperature profiles. In quantum mechanics, the equivalent function is given by the difference between the energy of the wavefunction and the height of the potential barrier, the latter being a known function of position. Here, the dependence of f on z is far more complicated since the behaviour of f does not depend linearly on $c^2(z)$ and consequently the behaviour of the potential is harder to predict.

The eigenvalues corresponding to equation (4.18) are (from Chapter 2):

$$\lambda_{1,2} = \frac{\gamma g}{2c^2} - \frac{k_x^2 c^{2\prime}}{2S} \pm \tfrac{1}{2}(\frac{k_x^4 (c^{2\prime})^2}{S^2} + \frac{2}{S} g k_x^2 c^{2\prime}(2\frac{k_x^2}{\omega^2} - \frac{\gamma}{c^2})$$
$$+ \frac{\gamma^2 g^2}{c^4} - 4\frac{(\gamma-1)g^2 k_x^2}{\omega^2 c^2} - 4\frac{S}{c^2})^{\frac{1}{2}}.$$

Written in terms of the acoustic cutoff and gravity-wave resonance frequencies this becomes:

$$\lambda_{1,2} = \frac{\omega_a}{c} - \frac{k_x^2 c^{2\prime}}{2S} \pm \tfrac{1}{2}\left(\frac{k_x^4 (c^{2\prime})^2}{S^2} + \frac{2}{S} g k_x^2 c^{2\prime} (2\frac{k_x^2}{\omega^2} - \frac{\gamma}{c^2}) + 4\frac{\omega_a^2}{c^2} - 4\frac{\omega_g^2}{\omega^2} k_x^2 - 4\frac{S}{c^2}\right)^{1/2} . \tag{4.20}$$

In the isothermal limit, the gradient terms vanish, the eigenvalues become ik_z as demonstrated in Chapter 2 and satisfy

$$-k_z^2 c - 2ik_z \omega_a + k_x^2 c \left(\frac{\omega_g^2}{\omega^2} - 1\right) + \frac{\omega^2}{c} = 0, \tag{4.21}$$

which can be rearranged into the form of the dispersion relation given by equation (4.11).

9.1.

Hines (1960) noted a difficulty with the formulation of the acoustic-gravity problem. Had we chosen a different dependent variable - the perturbed pressure, say - we would have obtained a different function for the potential leading to different reflection coefficients and implying that the physical variables of the wave are all reflected at different heights, assuming that reflection occurs at points where the potential vanishes (cf. Chapter 2). This ambiguity was addressed by Einaudi and Hines (1968) who examined the potential functions resulting from each choice of variable in turn. These authors determined that the WKBJ approximation broke down for all values of the wave parameters before any of the zeros of the potentials were encountered. They concluded that reflection was most likely to be indicated by the breakdown of the conditions for the WKBJ approximation and that, for $|f'/f|^2 \ll 1$ and $|f''/f| \ll 1$, the potentials were indistinguishable from one another. The only exception to this was found to be the isothermal potential, f_0, which differed markedly from the others whenever $\omega^2 \ll k_x^2 c^2$. (Notice that f_0 is the expression which is obtained by the local dispersion relation technique of making the coefficients of the dispersion relation functions of position. Figures (4.5a) and (4.5b) contrast the behaviour of the potential functions f_0 and f, given by equation (4.19), in the vicinity of $\omega = ck_x$ while Figure (4.5c) shows that f_0 and f vary similarly away from this neighbourhood.) Einaudi and Hines therefore concluded that the most appropriate potential to use in estimating the height at which reflection takes place is, (Einaudi and Hines equation [59]):

$$
\begin{aligned}
f_M &= \frac{\omega^2}{c^2} + k_x^2\left(\frac{\omega_g^2}{\omega^2} - 1\right) - \frac{\omega_a^2}{c^2} + \frac{g k_x^2}{\omega^2}\frac{c^{2\prime}}{c^2} \\
&= \frac{\omega^2}{c^2} + k_x^2\left(\frac{\omega_B^2}{\omega^2} - 1\right) - \frac{\omega_a^2}{c^2} \\
&= f_0 + \frac{g k_x^2}{\omega^2}\frac{c^{2\prime}}{c^2},
\end{aligned}
\tag{4.22}
$$

where ω_g and ω_B are the isothermal and non-isothermal Brunt-Väisälä frequencies respectively and f_M does not include higher order derivative terms which would arise from alternative variables and which are not significant if the WKBJ approximation applies. f_M is a logical extension of the reflection condition for waves in an isothermal atmosphere while incorporating some of the derivative terms. The intuitive prediction that reflection of atmospheric waves will occur near the non-isothermal Brunt-Väisälä frequency is supported, although this level can only be viewed as the height at which reflection must be at least suspected if not expected. It would appear that the different predicted reflection levels is a manifestation of the fact that reflection does not take place at a specific level but rather over a range of heights, being strongest near zeros of the potential where the parameter gradients are largest.

The theory of the solar 5-minute oscillation currently most favoured is that of acoustic modes trapped in the sub-photospheric convection zone. The upper reflecting surface is taken to be at the top of the convection zone while the lower level occurs where acoustic waves are reflected by the changing temperature as it increases towards the centre of the Sun. Essentially, reflection is to be anticipated whenever f_M passes through zero. By plotting the variation of ω_a and ω_B, through the assumed temperature variation with height in the solar atmosphere, Ando and Osaki (1977) predicted cavities where standing waves might exist for atmospheric waves. Observations of the oscillations with periods around 5 minutes agree well with the predictions of this model. Similarly, there is qualified agreement with the observations of 3-minute oscillations as acoustic waves trapped in a chromospheric cavity near the temperature minimum, although observational difficulties make this comparison more difficult. Even more controversial is the explanation of longer period waves as standing gravity waves trapped deep in the solar interior.

9.2.

Having discussed the height at which reflection takes place, we will now consider the degree of reflection. The variety of potentials is apparently more serious in this context since the reflection coefficient derived using phase integral methods (Heading, 1962), is dictated by the specific potential function, irrespective of the validity of the WKBJ solutions in the reflection region. We would anticipate that in the calculation of the coefficient of reflection, the differences between the various potentials would be compensated for by changes in the limits of integration between the choices of dependent variable.

As we stated earlier, we are seeking a system where the asymptotic solutions are plane waves (the derivative terms are small and the potential is positive) propagating through a central reflection region (large gradient terms and a negative potential) between two transition points. Unfortunately, none of the temperature profiles which are realistic descriptions, even of limited regions, of the solar or terrestrial atmospheric temperature structure lead to such a variation in the potential. Even in the special case of a vertically propagating wave ($k_x = 0$), the potential is not of this form for linear, parabolic or exponential temperature variation. In Figure (4.6), the potential given by equation (4.19) with $k_x = 0$ is plotted for the region of the solar temperature minimum where the temperature profile is approximately a quadratic function of height. It can be seen that this potential possesses most of the desired features but vanishes at infinity so that no waves can propagate there. We are thus unable to calculate analytically reflection coefficients and must look to numerical solutions of this problem.

10. Solution of the full differential equation

Before attempting to solve equation (4.18), we will combine the parameters in such a way as to produce a smaller, dimensionless set in order to write the differential equation in a more efficient form. First we transform the independent variable to $s = k_x z$ and define a the dimensionless quantity, $\hat{x} = c^2 k_x^2/\omega^2$. The remaining dimensional parameters may then be incorporated into $d_1 = g k_x/\omega^2$ and $d_2 = (\gamma - 1) d_1^2$, giving:

$$\hat{x}(1-\hat{x})\ddot{u}_z + (\dot{\hat{x}}\hat{x} - \gamma d_1(1-\hat{x}))\dot{u}_z + ((1-\hat{x})^2 + d_2(1-\hat{x}) - d_1\dot{\hat{x}}\hat{x})u_z = 0 \quad (4.23)$$

where $\dot{} = d/ds$.

We must now decide on the specific dependence of temperature, and hence sound speed, on height in order to complete the description of the problem. By choosing a linear variation such that $\hat{x} = \alpha s$ ($c^2 = \omega^2 \alpha z / k_x$), we are giving ourselves the best possible chance of deriving an analytic solution. This particular choice for $T(z)$ can only be an accurate description of the Earth's atmosphere over a limited range of heights (cf.§2.3) - the greatest flaw being the resultant zero sound speed at ground level, which is completely unphysical. Despite this highly simplistic model, we will find that the problem remains intractable in this form. Henceforth, we will therefore be considering the problem of acoustic-gravity wave propagation in an atmosphere with no wind shear and linear temperature variation with height.

Substituting the linear behaviour of $\hat{x}$ into equation (4.23) results in:

$$\hat{x}(1-\hat{x})\ddot{u}_z + (c_0\hat{x} - c_1)\dot{u}_z + (c_2\hat{x}^2 - c_3\hat{x} + c_4)u_z = 0 \tag{4.24}$$

where differentiation is now with respect to $\hat{x}$ and

$$c_0 = 1 + c_1 \quad c_1 = \frac{\gamma d_1}{\alpha} \quad c_2 = \frac{1}{\alpha^2} \quad c_3 = \frac{c_1}{\gamma} + c_2 + c_4 \quad c_4 = (\gamma - 1)\frac{c_1^2}{\gamma^2} + c_2$$

Since α and γ are known constants, equation (4.24) possesses two singularities corresponding to $c^2 = 0$ and $\omega^2 = k_x^2 c^2$. As was already noted, the singularity arising from the zero sound speed at the origin is completely unphysical and henceforth we will concentrate on the remaining "sonic line" singularity. To this end, we transform to a new coordinate system with the origin at the second singularity, $t = \hat{x} - 1$, giving (with respect to t)

$$t(t+1)\ddot{u}_z - (1 + c_0 t)\dot{u}_z + (f_1 + f_2 t - c_2 t^2)u_z = 0 \tag{4.25}$$

where $f_1 = c_1/\gamma$ and $f_2 = f_1 + (\gamma - 1)f_1^2$.

Before we attempt to find a general analytic solution of equation (4.25), we will use the position and nature of its singular points to classify it (Murphy, 1960). For the general second order ode,

$$y'' + p(x)y' + q(x)y = 0$$

the singularities are *regular* if and only if $p(x)$ is a pole of order 1 or less and $q(x)$ a pole of order 2 or less (Morse and Feshbach, 1953). For equation (4.25), which has singularities at $t = 0$ and $t = -1$,

$$p(t) = -\frac{1 + c_0 t}{t(1+t)} \qquad q(t) = \frac{f_1 + f_2 t - c_2 t^2}{t(1+t)}$$

so that both poles are simple. The final critical point is the point at infinity. In order to examine the behaviour of equation (4.25) at this point, we transform to $w = 1/t$ and examine the behaviour of the resulting equation at $w = 0$. The general ode in terms of w is:

$$\frac{d^2y}{dw^2}+P(w)\frac{dy}{dw}+Q(w)y=0,$$

where $P(w)=2/w-1/w^2p(1/w)$ and $Q(w)=1/w^4q(1/w)$. For $w=0$, the former is also a simple pole but, as the latter is a pole of order greater than 2, the point at infinity must be an irregular singular point.

The solutions of equations with regular singular points at $x=0,1$ and an irregular point at infinity are complicated functions. In some cases the equation may be transformed into the equation of the prolate or oblate spheroidal wavefunction. Alternatively, the equation may have the appropriate symmetries to permit transformation into the Mathieu equation:

$$\frac{d^2y}{dx^2}+(b-h^2\cos^2x)y=0$$

where b and h are arbitrary constants. Mathieu's equation has two irregular singular points (at 0 and ∞) and it can be shown that two independent solutions can be generated which possess a periodic dependence on position. This result, *Floquet's theorem* is discussed in more detail in Chapter 6 where the periodicity of the solutions is the keypoint.

Since there is no obvious transformation from equation (4.25) leading to a well-known equation with a standard solution technique it seems unlikely that this direction of investigation will be able to yield any useful results. However, by setting this same problem in terms of the dependent variable $\nabla\cdot\mathbf{u}\exp(-\frac{1}{2}\int(c^2/\gamma g)dz)$, Pitteway and Hines (1967) were able to derive an equation with only one singularity and possessing analytic solutions. The resulting equation was a form of Whittaker's confluent hypergeometric equation for which solutions are possible in terms of Laguerre functions or, in special cases, Bessel functions (Sneddon, 1956). These authors also derived solutions in terms of similar functions for an exponential temperature variation with height.

It seems that the limit of analytic treatment of this problem has been reached and for any more complicated temperature profile computer simulations would be required. In order to pursue an analytic approach, we will alter the direction of our investigation and so we will henceforth limit our region of interest to the neighbourhood of the $\omega^2=k_x^2c^2$ singularity.

11. Local solution

For an isothermal atmosphere, we noted in §5 that $\omega^2 = k_x^2 c^2$ represents Lamb waves which are evanescent and propagate horizontally. (This may be confirmed by substituting $\omega = k_x c$ into the definition of f_0 given by equation (4.22), showing that the isothermal potential is always negative for this combination of parameters). Since $u_z = 0$ for these waves, they can propagate in single modes along a solid, horizontal interface such as the surface of the Earth whereas other atmospheric waves would have to be coupled at such an interface as in a reflection process. It seems likely that the infinity introduced into equation (4.18) by the equality of ω^2 and $k_x^2 c^2$ might lead to a non-zero u_z for such waves in a non-isothermal atmosphere so that single modes cannot satisfy the boundary conditions at the rigid Earth's surface. Einaudi and Hines (1968) considered the region of parameter space surrounding this singularity to merit full wave analysis since the WKBJ approach becomes invalid in this neighbourhood. They suggested that in this region complex coupling of acoustic and gravity waves would occur and that the potential f_M given by equation (4.22) would therefore be inadequate to provide any information about the wave characteristics here.

In this section we will therefore focus our attention on the neighbourhood of the singularity where the sound speed resonates with the product of the frequency and the horizontal wavenumber ($t = 0$ in the notation of equation (4.25)). Near $t = 0$, terms of the order of t^2 will be negligible and a good approximation to equation (4.25) is given by:

$$t\,\ddot{u}_z - (1 + c_0 t)\,\dot{u}_z + (f_1 + f_2 t)\,u_z = 0. \tag{4.26}$$

A sequence of transformations is required to reduce equation (4.26) into a standard form for solution. First, let $u_z = \mathrm{v}\exp(kt)$ where k satisfies $k^2 - c_0 k + f_2 = 0$, reducing (4.26) to

$$t\,\ddot{\mathrm{v}} + (A_1 + B_1 t)\,\dot{\mathrm{v}} + A_2 \mathrm{v} = 0,$$

where $A_1 = -1$, $B_1 = 2k - c_0$ and $A_2 = f_1 - k$.

Now let $\mathrm{v}(t) = y(\hat{z})$ where $\hat{z} = -B_1 t$, giving

$$\hat{z}\,y'' + (A_1 - \hat{z})\,y' - \frac{A_2}{B_1}\,y = 0. \tag{4.27}$$

In general, equation (4.27) is the confluent hypergeometric equation (Bateman Manuscript Project Staff, 1953). On evaluating A_1, A_2 and B_1, we find that we have a special case of this equation and need not resort to solution via confluent hypergeometric series. First, we solve for k in terms of our original coefficients.

The equation is a quadratic with solutions given by

$$2k = c_0 \pm \sqrt{c_0^2 - 4f_2}$$

$$= 1 + c_1 \pm \left((1+c_1)^2 - 4\frac{c_1}{\gamma}\left(1 + (\gamma-1)\frac{c_1}{\gamma}\right)\right)^{1/2}$$

$$= 2\left(1 + c_1\left(1 - \frac{1}{\gamma}\right)\right) \quad \text{or} \quad 2\frac{c_1}{\gamma}.$$

Hence

$$\frac{A_2}{B_1} = \frac{c_1/\gamma - (1+c_1-c_1/\gamma)}{2(1+c_1-c_1/\gamma) - (1+c_1)} = -1 \quad \text{or} \quad \frac{A_2}{B_1} = \frac{f_1 - k}{2k - c_0} = \frac{c_1/\gamma - c_1/\gamma}{2c_1/\gamma - (1+c_1)} = 0.$$

Thus, we must solve either

$$\hat{z}\, y'' - (1+\hat{z})y' + y = 0 \quad \text{or} \quad \hat{z}\, y'' - (1+\hat{z})y' = 0.$$

We solve the former (but the solutions of the latter reproduce these) to find:

$$y = A \exp(\hat{z}) + B(1+\hat{z}),$$

where A, B are arbitrary constants. Transforming back to the original variables yields:

$$u_z = A \exp\left(\frac{c_1}{\gamma}t\right) + B\left(1 + \left(2\frac{c_1}{\gamma} - c_1 - 1\right)t\right) \exp\left(\left(1+c_1-\frac{c_1}{\gamma}\right)t\right).$$

These solutions of equation (4.26) are well-behaved in the neighbourhood of $t=0$ and do not become singular. Setting the problem in terms of one of the alternative variable combinations mentioned earlier ($\rho_0^{1/2}\nabla\cdot\mathbf{u}$, say) is therefore highly recommended since the complication of this additional singularity may thus be avoided.

12. Summary

We have seen that acoustic-gravity waves propagating in inhomogeneous atmospheres pose many intriguing problems. Unfortunately, all attempts at analytic solution are thwarted at every turn by the unyielding nature of the equations. The only analytic solutions possible for the case of an atmosphere with a non-zero temperature gradient occur for either linear or exponential temperature profiles, neither of which gives a realistic description of either the solar or terrestrial atmospheres over more than a short range of heights. (It was also shown that the possibility of deriving these solutions is sensitive to the original choice of field variable.) We were also unable to calculate the degree of reflection analytically for similar reasons.

Without an analytic formula for the reflection coefficient, we must appeal to a diagrammatic method in order to determine if significant reflection is to be anticipated. In Chapter 2, a point where the eigenvalues undergo a close approach without crossing was interpreted as being likely to indicate mode conversion (Diver, Ph.D. thesis, 1986). We can see from Figure (4.7) that the eigenvalues do not behave in this way and so we postulate that mode conversion will be small.

To increase the physical realism of the model, we should also include the effects of wind shear - mathematically, a nonzero u_{0x} - in the equations of motion but such an attempt simply renders the differential equation describing the system even more unwieldy. Pitteway and Hines (1967) are amongst several authors who have attempted this step but, in order to make any headway, they found that it was necessary to ignore the effect of the background wind and it was only under the assumption of vanishing $\mathbf{u}_0$ that they were able to derive solutions for the special cases of temperature variation described above.

We will now progress to a problem in plasma physics where there are two truly distinct modes of oscillation. This scenario, involving two different types of wave, offers far greater scope for future investigations of linear mode conversion than atmospheric waves. As an introduction, we will consider the propagation of waves in a homogeneous cold plasma in the following chapter, much as we considered waves in an isothermal atmosphere initially in this chapter.

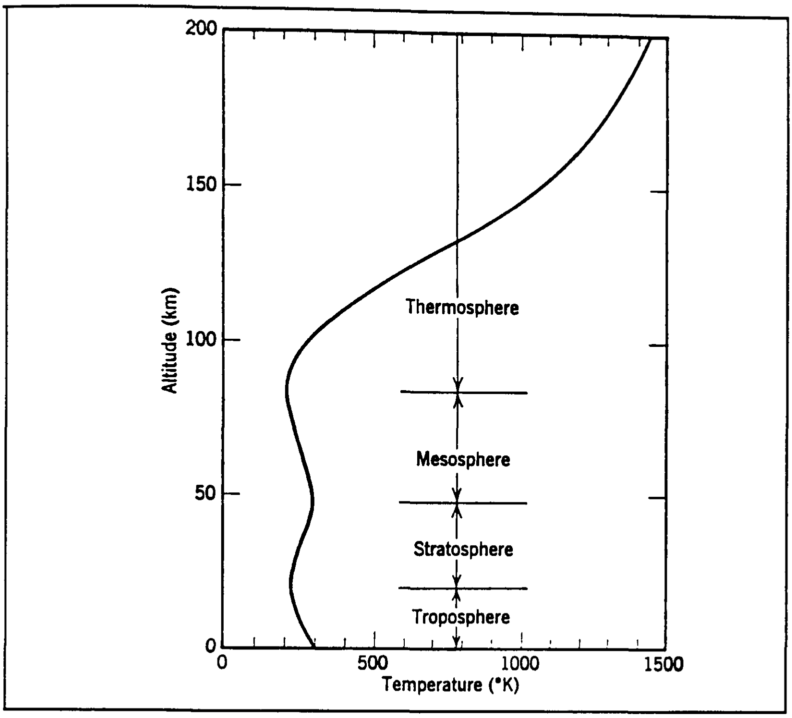


Figure (4.1) Typical temperature variation with height in the Earth's atmosphere.

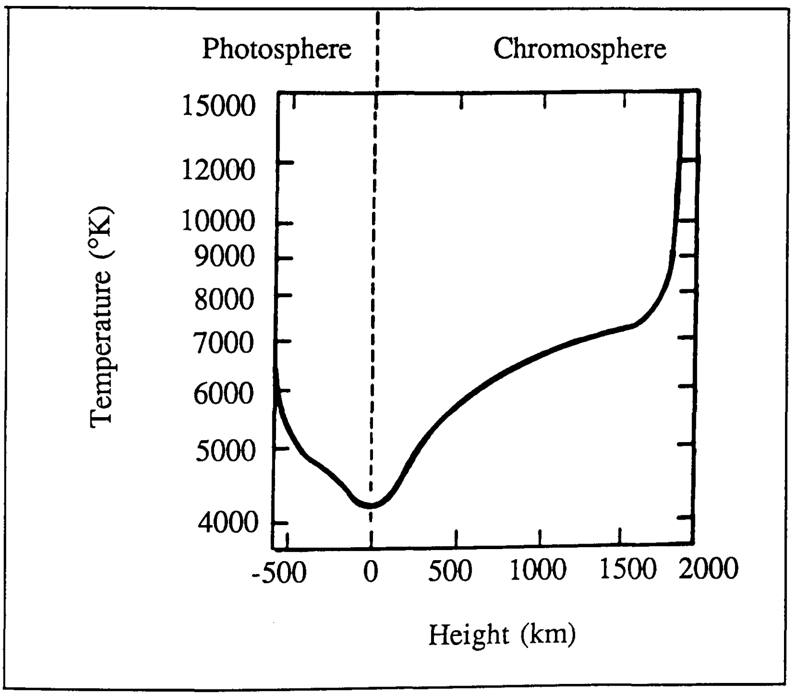


Figure (4.2) Temperature variation with height in the Sun's *photosphere* and *chromosphere* showing the temperature minimum (dashed line).

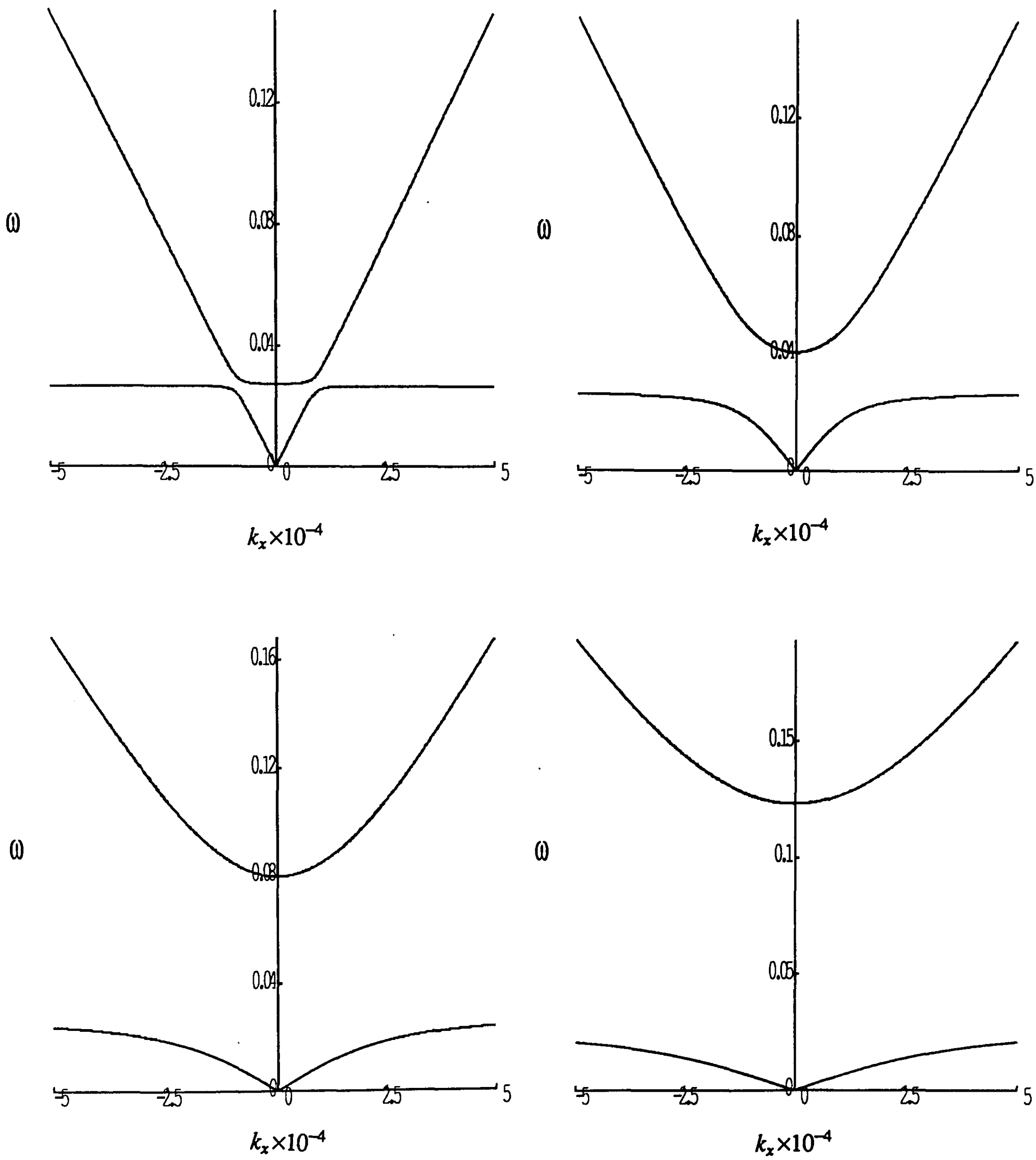


Figure (4.3) Dispersion relation curves, ω (s) versus k_x(m^{-1}) for a range of values of k_z(m^{-1}). Clockwise from top left:- $k_z = 0$, $k_z = 10^{-4}$, $k_z = 2.5\times10^{-4}$, $k_z = 4\times10^{-4}$. ($c = 300$ms^{-1}.)

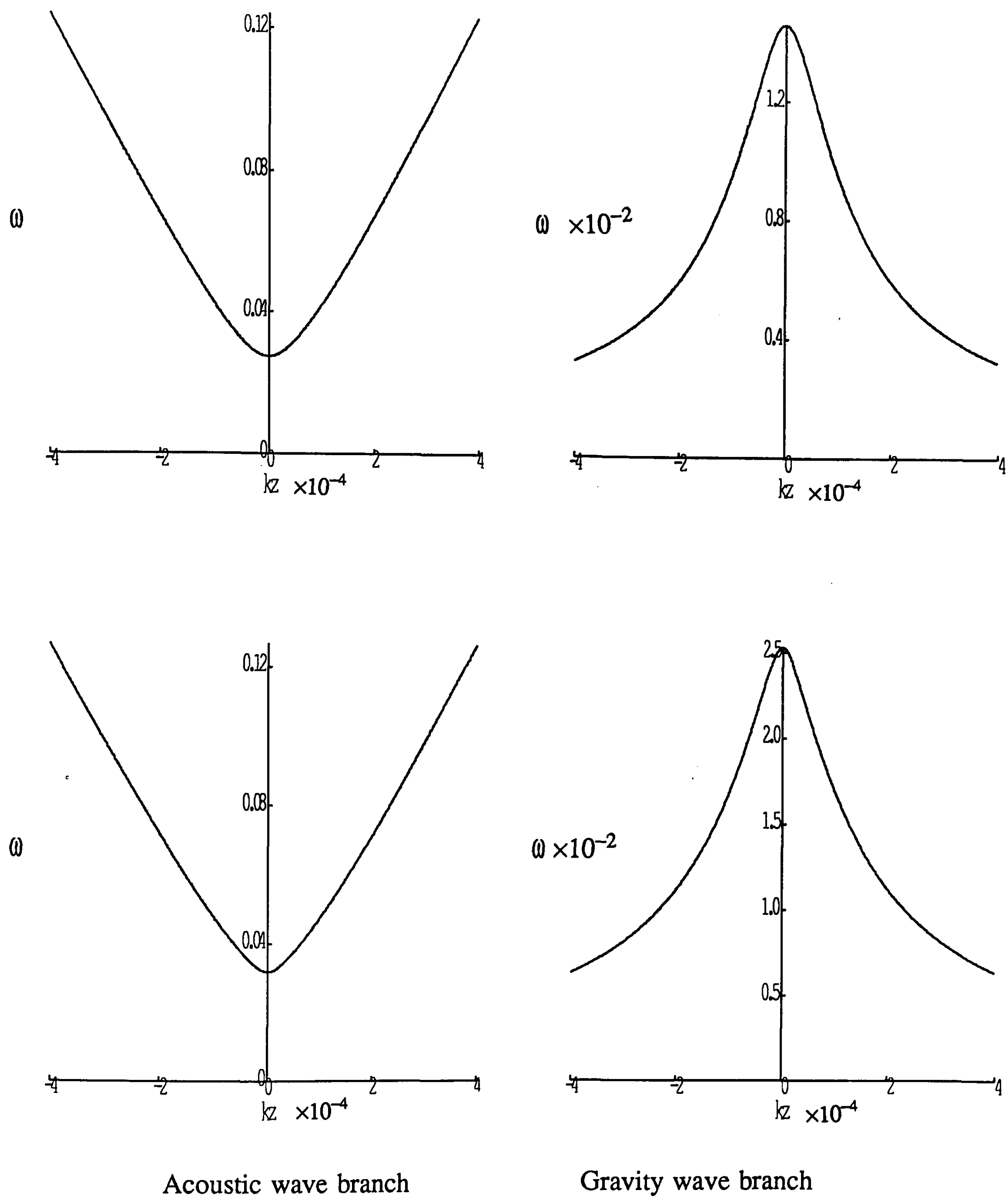
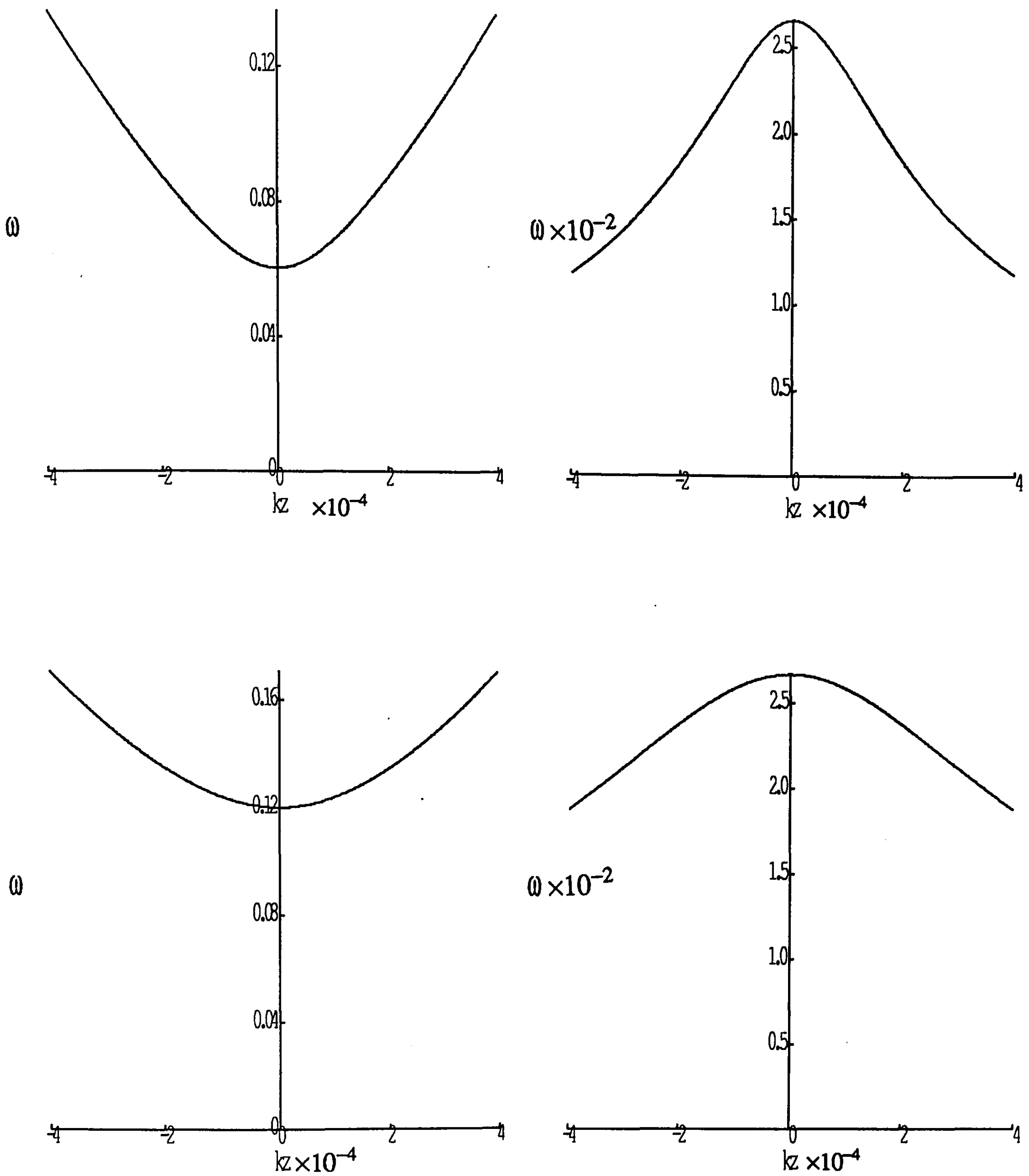


Figure (4.4a) A set of dispersion relation curves, ω versus k_z for different k_x. Note that for vertical propagation ($k_x = 0$) only acoustic waves may propagate and the gravity wave curve will vanish. Top:- $k_x = 5\times10^{-5}$. Bottom:- $k_x = 10^{-4}$. ($c = 300\text{ms}^{-1}$.)



Acoustic wave branch　　Gravity wave branch

Figure (4.4b) Plot of ω versus k_z for k_x:- Top, $k_x = 2\times10^{-4}$. Bottom, $k_x = 4\times10^{-4}$.

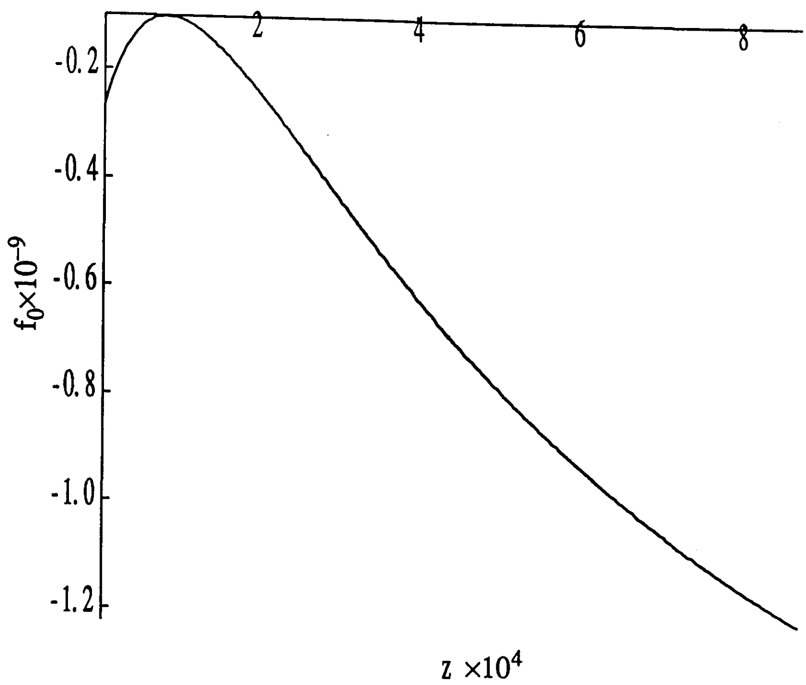


Figure (4.5a) Plot of the isothermal potential function (cf. equation (4.22)) for a linear temperature profile, $T = A + Bz$ where $A = 270$ °K and $B = 0.01$ °K m^{-1} in the neighbourhood of $\omega = ck_x$ for $\omega = 2\times10^{-2}s^{-1}$ and $k_x = 5\times10^{-5}m^{-1}$.

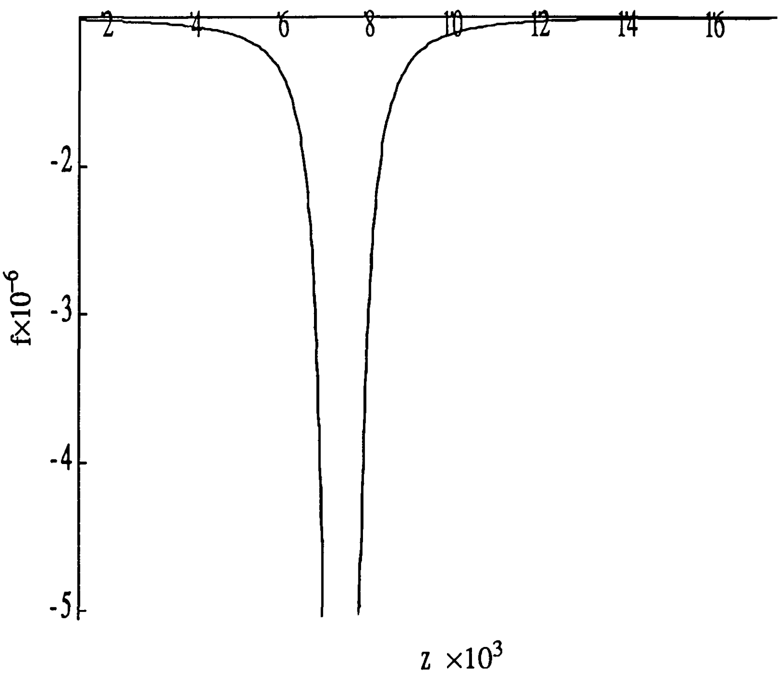
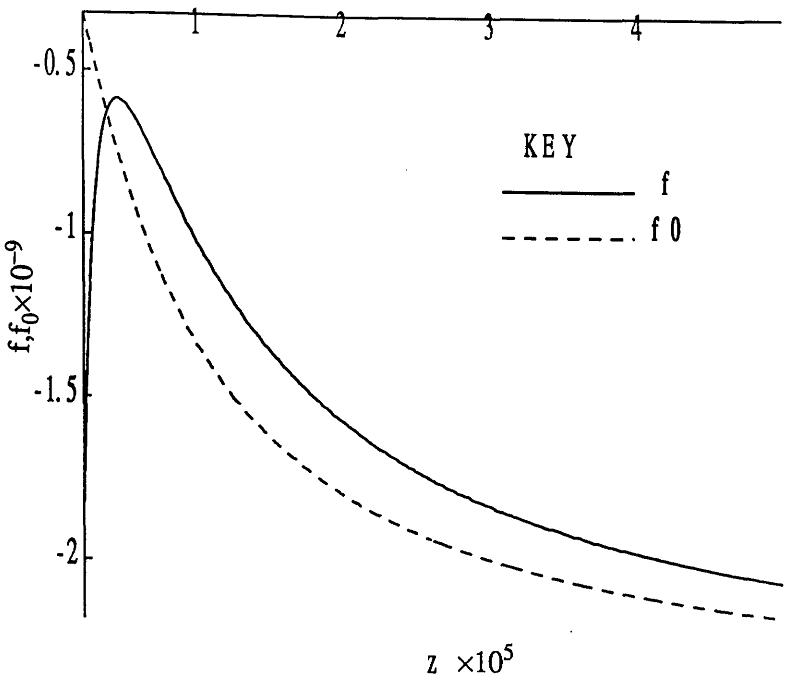


Figure (4.5b) Plot of the potential function given by equation (4.19) centred round the singularity where $\omega = ck_x$. (The temperature profile, frequency and horizontal wavenumber are the same as above.) The singularity occurs at the height $z = 7495$ m.



Figure(4.5c) Comparison of the height variation of the isothermal potential and the potential given by equation (4.19) for the temperature profile and parameters of Figures (4.5a and b) It can be seen that at heights away from $\omega = ck_x$ the two potentials are very similar.

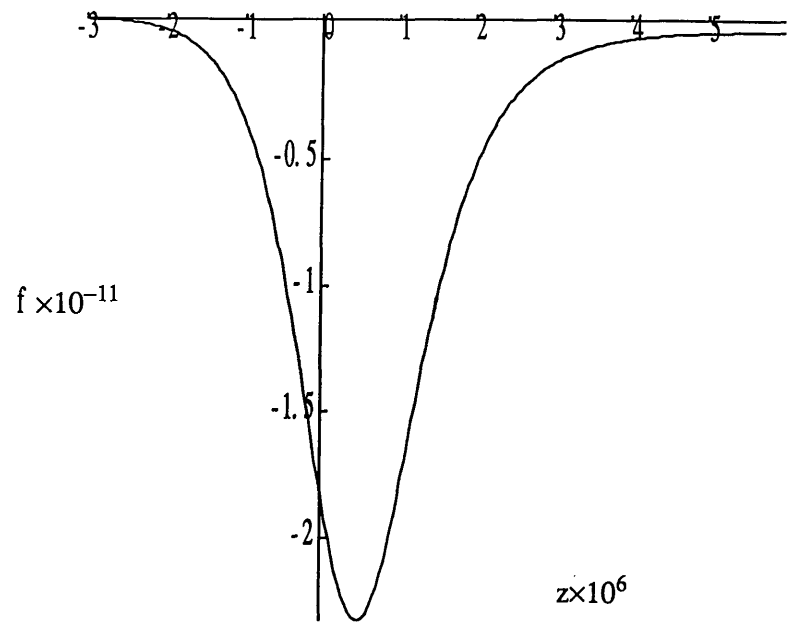


Figure (4.6) Variation of the potential function given by equation (4.19) with height for a parabolic temperature profile representing the neighbourhood of the Sun's temperature minimum. $T = A + Bz + Cz^2$ where A=4.3×10^{-3} °K, $B = -1.4\times10^{-3}$ °K m^{-1}, $C = 1.8\times10^{-9}$ °K m^{-2}. $\omega = 1.5\times10^{-2}$s^{-1}, $k_x = 0$. (Note that for the Sun, $g = 274$ms^{-2} and $M = 1.3$ kg/kmol.)

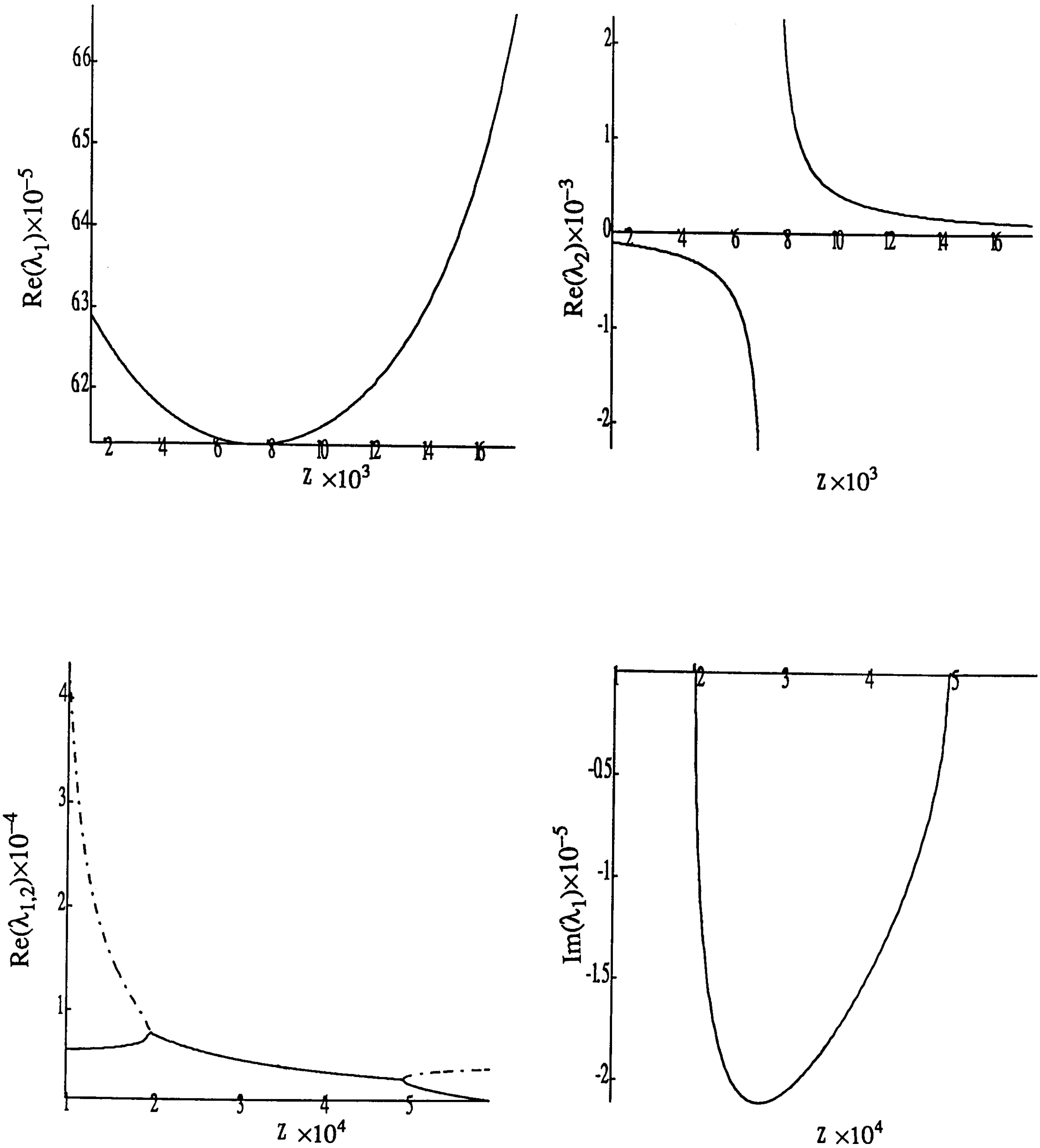


Figure (4.7) Eigenvalues, $\lambda_{1,2}$ versus z for the same parameters as Figure (4.5). Top:- Near the singularity $\omega = ck_x$ one eigenvalue becomes singular while the other remains finite. Bottom:- Beyond the singularity, when the imaginary parts of the eigenvalues become non-zero ($\text{Im}(\lambda_1) = -\text{Im}(\lambda_2)$), the real parts become equal.

Chapter 5 - Review of wave propagation in a cold plasma with a uniform magnetic field

1. Introduction

The culmination of our discussions of waves in inhomogeneous media will come in the following chapter when we derive the complete, analytic solution for waves in a cold plasma where the magnetic field is no longer uniform but rotates in space about an axis. In order to gain insight into the wave motions under such conditions, we must first consider the much more straightforward case of waves propagating in a cold plasma where the magnetic field is invariant.

The cold plasma model was initially discussed in Chapter 3 where its governing equations were derived from Kinetic Theory. To recap, the basic assumption of this model is that the random motions of the plasma particles are negligible so that the elements of the pressure and heat tensors, which result from these random thermal motions, may be disregarded. The zero temperature assumption means that the velocity distribution function for a cold plasma species is therefore a Dirac delta function centred on the macroscopic flow velocity: $f_s(\mathbf{r},\mathbf{v},t) = \delta[\mathbf{v} - \mathbf{u}_s(\mathbf{r},t)]$. Although this is a highly idealised model, it has proved popular in the study of plasma waves for several reasons. First, all possible wave modes for this model are derivable. Because a magnetised plasma possesses a bewildering number of oscillatory modes, it is a great advantage to consider, as a first approximation, a simplified system in which all the modes may be calculated and the basic physics understood. The cold plasma model also provides a surprisingly comprehensive view of plasma waves and even an accurate description of the common small-amplitude perturbations of hot plasmas. In addition, this model is aesthetically pleasing since it is self-consistent, an attribute which will prove crucial to the arguments of Chapter 6. Many plasma waves which have been observed experimentally can be explained in terms of either of the two basic cold plasma modes, which is a strong testimony on its behalf. Recent developments in the field of solid state plasmas, which are of fundamental importance in the production of computer chips, and which are probably the plasmas which most closely resemble the conditions of a cold plasma, make the study of such a model particularly relevant to this branch of technological physics.

Further simplifying assumptions which we will employ to make the problem of wave propagation in a cold plasma more tractable are that there is no dissipation, that terms quadratic in the perturbed quantities are negligible and that the plasma is infinite in extent. We are therefore excluded from investigating collisional effects, non-linear phenomena such as shocks, and boundary effects. In general, collisions result in the wave motions being damped and so neglecting collisional processes (i.e. dropping the term $\mathbf{A}_s$ from the momentum equation, (3.5)) may be regarded as omitting damping from our description of the plasma. The plane wave solutions used in Chapter 4 will again be used but this is not restrictive since all physically reasonable wave motions may be reconstructed from a superposition of such plane waves. A final assumption which is often made is to neglect the ion motion compared to that of the electrons because their greater mass means they form a fixed background at higher frequencies. This is not an essential constraint and retaining the complete, two-species description simply involves summation signs. Whenever this last assumption is invoked, it will be stated explicitly.

There are two principle methods of solving the cold plasma equations. The first is a straightforward simultaneous solution of Maxwell's equations and the equations for the particle motions, while the second, which we favour, involves the construction of intermediate relations between the field variables, especially between the current and electric field. In the latter approach, the plasma is considered to be analogous to a dielectric material possessing internal currents. The current is related to the electric field by $\mathbf{J} = \sigma\mathbf{E}$, where σ is the conductivity tensor, and hence the electric displacement is similarly related to the electric field via the dielectric tensor. The common aim of both methods is the construction of a dispersion relation which is the key to any homogeneous plasma mode since it contains all of the information about the propagation of the mode except its phase.

The properties of plasma waves are retained over a wide range of parameters and a classification scheme which exploits this is the CMA diagram, named after its authors P.C. Clemmow, R.F. Mullaly and W.P. Allis. The axes of the CMA diagram reflect the plasma parameters - electron density, magnetic field strength, percentage composition by ion species and wave frequency. For the two-component plasma under discussion, two axes suffice to incorporate all this information and the CMA diagram is a two-dimensional graph with the boundaries between regions of different wave characteristics being marked by simple curves. As shown in Figure (5.1), the abscissa represents $(\omega_{pe}^2 + \omega_{pi}^2)/\omega^2$ where ω is the wave frequency, so that the density increases to the right or, for fixed density, the frequency decreases. The ordinate represents $\Omega_e \Omega_i/\omega^2$ and the ambient magnetic field, $\mathbf{B}_0$ increases upwards or, for fixed $\mathbf{B}_0$, the frequency decreases. Thus, keeping the plasma parameters fixed, the frequency decreases radially away from the origin. The characteristic shared by all the waves in a particular region of the CMA diagram is the topological form of the *wave-normal surface*, the locus of the tip of the vector having the direction of wave propagation and the magnitude the wave velocity divided by the speed of light - $\omega \mathbf{k}/ck^2$. In terms of the refractive index, $\mathbf{n} = c\,\mathbf{k}/\omega$, the wave normal surface is the locus of the tip of its reciprocal vector, $\mathbf{n}^{-1} \equiv \mathbf{n}/n^2$. Thus the CMA diagram is a polar plot of the normalised phase velocity, v_{ph}/c, drawn in parameter space. Once we have examined the derivation of the dispersion relation from the dielectric description of the cold plasma, we will discover the equations for the lines bounding the regions of the CMA diagram and will discuss the assignment of particular wave modes to their appropriate sections.

2. Constructing the dielectric tensor

The equations of the cold plasma model are the "two-fluid" equations under the approximation of a cold plasma plus Maxwell's equations (cf. Chapter 3). To avoid copious repetition, we restate here only the linearised forms of the equations which we will currently require in order to derive a single differential-equation description of this model. These are, respectively, the equation of motion, Maxwell's two curl equations and the equation relating current to velocity:

$$m_s \frac{\partial \mathbf{v}_s}{\partial t} = q_s(\mathbf{E} + \mathbf{v}_s \times \mathbf{B}_0) \,, \tag{5.1}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \,, \tag{5.2}$$

$$\nabla \times \mathbf{B} = \mu_0 \, (\mathbf{J} + \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}) \,, \tag{5.3}$$

$$\mathbf{J} = n_0 \sum_s q_s \, \mathbf{v}_s \,, \tag{5.4}$$

where the sum is over all particle species, s, and where we have used the convention that vector quantities which are not subscripted zero are first order. The plasma is neutral overall and, assuming for simplicity that the ions are singly charged, we must have $n_e = n_i$ in the equilibrium state, which we have denoted by n_0. Henceforth, we will also assume that the plasma ions are simply hydrogen, so that the charges on the two species present are equal and opposite - extending this to different ions would only require multiplication of the ion charge by a factor of the atomic number. The notation $q_s = \varepsilon_s e$ will be used throughout to represent the charge on a species, where e is the electronic charge and $\varepsilon_s = \begin{cases} -1 & \text{for electrons} \\ +1 & \text{for ions} \end{cases}$.

In the manipulations which follow, we adopt the approach of other authors, for example Stix (1962) and Chen (1984). Considering the plasma to be a dielectric material, we may restyle equation (5.3) in the form:

$$\nabla \times \mathbf{B} = \mu_0 \dot{\mathbf{D}},$$

where **D** is the electric displacement, which includes both the vacuum displacement and the plasma current, and we use $\dot{} \equiv \partial/\partial t$. We aim to write $\mathbf{D} = \varepsilon \cdot \mathbf{E}$, where ε is the dielectric tensor replacing the dielectric constant of a simple dielectric material. By expressing the relationship between the plasma current and the electric field in terms of a conductivity tensor, $\mathbf{J} = \sigma \cdot \mathbf{E}$, we are assuming that the current is directly proportional to the electric field but not necessarily aligned along it. The anisotropy in the dielectric tensor arises from the presence of the externally applied magnetic field, $B_0 \hat{\mathbf{z}}$. Assuming harmonic plane wave solutions, we Fourier transform in the usual manner and, substituting from equation (5.3), we find:

$$\mu_0 (\mathbf{J} - i\omega\varepsilon_0 \mathbf{E}) = -i\omega \mathbf{D} \mu_0,$$

which defines the electric displacement, in terms of the conductivity tensor, to be:

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \frac{i\sigma}{\omega}\mathbf{E} = \varepsilon_0 (\mathbf{I} + \frac{i\sigma}{\varepsilon_0 \omega})\mathbf{E} = \varepsilon_0 \varepsilon \cdot \mathbf{E}.$$

To calculate σ from the relation between the current and the electric field, we see from equation (5.4) that we must first derive an expression for the velocity components in terms of the electric field. This may be done by rearranging equation (5.1) to get **v**(**E**). The components of the velocity may then be written:

$$v_{sx} = \frac{iq_s (E_x + i\varepsilon_s \Omega_s/\omega E_y)}{m_s \omega (1 - \Omega_s^2/\omega^2)}$$

$$v_{sy} = \frac{iq_s (E_y - i\varepsilon_s \Omega_s/\omega E_x)}{m_s \omega (1 - \Omega_s^2/\omega^2)} \qquad v_{sz} = \frac{iq_s}{m_s \omega} E_z,$$

where we have used the definitions of the cyclotron and plasma frequencies of a species, $\Omega_s = |q_s B_0|/m_s$ and $\omega_{ps}{}^2 = n_s q_s{}^2/\varepsilon_0 m_s$, respectively. These quantities will thus be prominent in the definitions of the elements of the dielectric tensor, which we are about to derive. In order to eliminate all first-order variables from the set of equations (5.1) to (5.4) in favour of the perturbed electric field, we substitute the above expressions for the velocity components into the current density equation, (5.4). This last relation is best expressed via the conductivity tensor and, from this, we can construct the dielectric tensor using the definition above:

$$\varepsilon = \begin{bmatrix} S & -iD & 0 \\ iD & S & 0 \\ 0 & 0 & P \end{bmatrix},$$

where

$$P = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2}$$

$$R = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2}\frac{\omega}{\omega+\varepsilon_s\Omega_s} \qquad L = 1 - \sum_s \frac{\omega_{ps}^2}{\omega^2}\frac{\omega}{\omega-\varepsilon_s\Omega_s},$$

$$S = ½(R+L) = 1-\sum_s \frac{\omega_{ps}^2}{\omega^2-\Omega_s^2} \quad , \; D = ½(R-L) = \sum_s \frac{\omega_{ps}^2\varepsilon_s\Omega_s}{\omega(\omega^2-\Omega_s^2)}.$$

3. The wave equation, the refractive index and the dispersion relation

The desired wave equation describing electromagnetic wave propagation may now be found by taking the "curl" of the equation (5.2):

$$\begin{aligned}\nabla\times(\nabla\times \mathbf{E}) &= -\nabla\times\dot{\mathbf{B}} \\ &= -\mu_0\ddot{\mathbf{D}} \\ &= -\mu_0\varepsilon_0\varepsilon\cdot\ddot{\mathbf{E}}.\end{aligned}$$

We may thus write the wave equation in its final form after Fourier transforming in time as:

$$\nabla\times\nabla\times\mathbf{E} - \frac{\omega^2}{c^2}\varepsilon\cdot\mathbf{E} = 0. \tag{5.5}$$

The coefficients of equation (5.5) are constant throughout space and so we Fourier transform in all space dimensions. Letting θ be the angle which the wavevector, **k**, makes with the magnetic field, and choosing the coordinate axes so that **k** has no component in the y−direction, the wave equation takes the form:

$$\mathbf{n}\times(\mathbf{n}\times\mathbf{E}) + \varepsilon\cdot\mathbf{E} = 0,$$

which is best expressed in matrix notation as $\mathbf{K}\cdot\mathbf{E} = 0$ where

$$\mathbf{K} = \begin{bmatrix} S-n^2\cos\theta^2 & -iD & n^2\sin\theta\cos\theta \\ iD & S-n^2 & 0 \\ n^2\sin\theta\cos\theta & 0 & P-n^2\sin^2\theta \end{bmatrix}.$$

The condition for obtaining a non-trivial solution of such a set of three simultaneous, homogeneous equations is that the determinant of the square matrix of coefficients, $|\mathbf{K}|$, must vanish. This condition generates the dispersion relation relating the frequency and wavenumber which we choose to express in terms of the refractive index vector, **n**. This dispersion relation has been obtained by Åström (1950), Sitenko and Stepanov (1956) and Allis (1959):

$$An^4 - Bn^2 + C = 0, \tag{5.6}$$

where $A = S\sin^2\theta + P\cos^2\theta$,

$$B = RL\sin^2\theta + PS(1+\cos^2\theta),$$

$$C = PRL.$$

(Note that we have made use of the identity $S^2 - D^2 = RL$, cf. §2.) Since equation (5.6) is a quadratic in n^2, it has two solutions which are in general labelled the *fast* and *slow* modes but for particular orientations and in particular frequency ranges, they may have specific local names. For a given frequency, there are only two types of wave which can propagate with their own distinct properties, corresponding to each solution of equation (5.6) and the two values of n which give the same n^2 simply represent forward and backward propagation of the same wave.

An alternative form of the dispersion relation, derived by Åström and Allis, may be obtained from equation (5.6) and is particularly useful when examining the dependence of propagation on the angle θ. We use the basic trigonometric identity, $\sin^2\theta + \cos^2\theta = 1$, and the definitions of A, B and C to rewrite equation (5.6) as:

$$n^4(S\sin^2\theta + P\cos^2\theta) - n^2(RL\sin^2\theta + 2PS\cos^2\theta + PS\sin^2\theta)$$
$$+ PRL(\sin^2\theta + \cos^2\theta) = 0,$$

which may be rearranged to give:

$$\sin^2\theta\,(Sn^4 - (PS + RL)n^2 + PRL) = -P\cos^2\theta\,(n^4 - 2Sn^2 + RL).$$

Finally, we use the definition of S plus a little more manipulation to obtain Åström and Allis's form of the dispersion relation:

$$\tan^2\theta = \frac{-P(n^2-R)(n^2-L)}{(Sn^2-RL)(n^2-P)}. \tag{5.7}$$

Before progressing to an analysis of equation (5.7) for specific angles, θ, we introduce the standard nomenclature used to classify waves in plasmas. The words *parallel* and *perpendicular* refer to the angle between the direction of propagation and the ambient magnetic field whereas *transverse* and *longitudinal* describe the angle between the wavevector and the oscillating electric field. Also, if the perturbed part of the magnetic field is zero, the wave is dubbed *electrostatic*, otherwise it is *electromagnetic*. Thus a wave which is longitudinal satisfies $\mathbf{k}\times\mathbf{E} = 0$ and so, by Faraday's Law, equation (5.2), $-i\,\omega\mathbf{B}_1 = 0$ so that the wave is also electrostatic. Conversely, a transverse wave will always be electromagnetic.

4. Parallel and perpendicular propagation

In this section, we investigate the dispersion relations for the special cases of parallel and perpendicular propagation which may be quickly obtained from equation (5.7). Propagation at any other angle involves much more complicated formulae because, as we can see from the matrix **K**, the E_x and E_y components only decouple from E_z at the principal angles $\theta = 0°$ and $\theta = 90°$. At these angles, the solutions take particularly simple forms and so a thorough examination of the nature of the modes in these special cases, which also provide limiting cases of the general propagation properties, is warranted.

4.1. Parallel propagation

By the definitions of parallel and perpendicular given above, the propagation is parallel if the angle between the wavevector and the background magnetic field is zero. On inserting $\theta = 0°$, into equation (5.7), we find that solutions are:

$$P = 0 \qquad\qquad n^2 = R \qquad\qquad n^2 = L.$$

The first of these solutions, $P = 0$, represents the longitudinal electron plasma oscillations introduced in Chapter 1 plus a small correction factor due to the presence of the ions. This is easily demonstrated by noting that, for $m_e/m_i \ll 1$, $P = 0$ is equivalent to $\omega^2 \approx \omega_{pe}^2$. Thus, $P = 0$ represents plasma oscillations which are unaffected by the magnetic field. In the cold plasma approximation, these oscillations do not propagate and so do not constitute one of the modes of the plasma. If the thermal velocity were not neglected, however, the dispersion relation for these oscillations would become $\omega^2 = \omega_{pe}^2 + k^2 V_e^2$ (where V_e is the electron thermal velocity). This is a travelling wave since its group velocity is not zero and represents one of the additional modes of oscillation which is possible for a "warm" plasma. Physically, the electron's thermal velocity causes regions neighbouring the oscillation to be affected by it and so information is carried out of the oscillating layer. (In practice, fringing effects of the electric field in any finite plasma would also couple the oscillation into other plasma regions and the assumption that Langmuir oscillations do not propagate even in a cold plasma is a highly idealised one.)

The other two modes for propagation parallel to the background magnetic field in a cold plasma, $n^2 = R$ and $n^2 = L$, are dependent on the wavenumber and are therefore the two true oscillatory modes. The solutions R and L are differentiated by their polarisations.

The polarisation of an electromagnetic wave is obtained from the polarisation of the electric field vector, since the magnetic induction vector may always be calculated from this using equation (5.2). The polarisation of the electric field vector is calculated from $\mathbf{E} = (E_1\hat{\mathbf{e}}_1 + E_2\hat{\mathbf{e}}_2)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, where $\mathbf{E}_{1,2}$ are two linearly independent, linearly polarised waves aligned along the unit vectors $\hat{\mathbf{e}}_{1,2}$ with complex amplitudes $E_{1,2}$. A complex amplitude may be considered to be the product of a real amplitude and a phase factor and, since it is the phase difference between the two waves which is important, we rewrite the full wave as:

$$\mathbf{E} = (E_{10}\hat{\mathbf{e}}_1 + E_{20}e^{i\delta}\hat{\mathbf{e}}_2)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)},$$

where E_{j0} (j=1,2) represents the real wave amplitude and δ the phase difference. For the particular case of a phase difference of 90°, $\delta = \pm\pi/2$, and we have

$$\mathbf{E} = (E_{10}\hat{\mathbf{e}}_1 \pm iE_{20}\hat{\mathbf{e}}_2)e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}, \tag{5.8}$$

where, for convenience, we assume that the polarisation vectors are aligned along the x and y axes and that the wave is propagating in the z direction which allows us to write:

$$E_x(z,t) = E_{x0}\cos(kz - \omega t) \ , \ E_y(z,t) = \mp E_{y0}\sin(kz - \omega t).$$

These expressions enable us to demonstrate that such a wave is in general elliptically polarised:

$$\frac{E_x^2}{E_{x0}^2} + \frac{E_y^2}{E_{y0}^2} = 1.$$

Two special cases which are of interest are *linear polarisation*, when either E_{x0} or E_{y0} is zero, and *circular polarisation* which occurs whenever $E_{x0} = E_{y0}$. If, looking along the positive z-axis, the wave is observed to be rotating clockwise, we have right polarisation which corresponds to the upper sign in equation (5.8). An outgoing wavefront rotating in the opposite sense is left polarised. Thus the components of circularly polarised waves have the same magnitude but are $\pi/2$ out of phase and will be indicated by complex amplitudes which satisfy $iE_x/E_y = \pm 1$.

The expression for the polarisation of an oblique wave in a cold plasma may be found from the second row of the matrix **K**:

$$\frac{iE_x}{E_y} = \frac{n^2 - S}{D},$$

so that the polarisation of the R mode is given by:

$$\frac{iE_x}{E_y} = \frac{R-S}{D} = 1,$$

and R is therefore *right circularly polarised* (RCP). For the L mode,

$$\frac{iE_x}{E_y} = \frac{L-S}{D} = -1,$$

which makes it *left circularly polarised* (LCP).

The plane of polarisation of an electromagnetic wave travelling along the line of a magnetic field will be rotated as it progresses, a phenomenon known as Faraday rotation. This effect is comparatively small and is only useful as a diagnostic tool under such extreme circumstances as the relatively large magnetic fields or very long path lengths encountered in interstellar space.

4.2. Perpendicular propagation

We will study perpendicular propagation in greater depth than the parallel case, since we will ultimately wish to describe perpendicular propagation in an *inhomogeneous* plasma and therefore desire a comprehensive and detailed homogeneous model with which to compare it. For $\theta = 90°$, the solutions of equation (5.7) are:

$$n^2 = \frac{RL}{S} \quad \text{and} \quad n^2 = P.$$

The polarisations of the RL/S and P modes are, respectively:

$$\frac{iE_x}{E_y} = \frac{P-S}{D} \quad \text{and} \quad \frac{iE_x}{E_y} = \frac{RL/S-S}{D} = -\frac{D}{S}.$$

Thus both of these modes are in general elliptically polarised.

In addition to considering the polarisation of the perpendicular modes, we will also study the variation of the refractive index with frequency which will again be relevant to the latter parts of Chapter 6, but first we introduce a new notation to facilitate further analysis.

5. The Appleton-Hartree equation

It is suitable to define the required new notation in a short digression on the Appleton-Hartree equation. This well-known equation (see for example Stix (1962)), which was first derived by Hartree (1931) and Appleton (1932), describes oscillations at sufficiently high frequencies that the motion of the ions may be neglected and so deals exclusively with electron modes in the plasma. It has been used with considerable success to study radio wave propagation in the ionosphere under the influence of the Earth's magnetic field (Budden (1961)).

Prior to deriving the Appleton-Hartree equation, we introduce some simplifying notation involving the dimensionless variables,

$$X = \frac{\omega}{\omega_{pe}} \quad \text{and} \quad Y = \frac{\Omega_e}{\omega_{pe}},$$

where the subscript e indicate that these are the plasma and cyclotron frequencies for electrons. These are simply related to the ion frequencies via the ratio of the electron to ion masses, $r = m_e/m_i$. In terms of these new variables, the basic elements of the dielectric tensor are:

$$P = 1-(\frac{\omega_{pe}^2}{\omega^2}+\frac{\omega_{pi}^2}{\omega^2}) = 1-\frac{1+r}{X^2},$$

$$R = 1-(\frac{\omega_{pe}^2}{\omega^2}\frac{\omega}{\omega-\Omega_e}+\frac{\omega_{pi}^2}{\omega^2}\frac{\omega}{\omega+\Omega_i}) = 1-(\frac{1}{X(X-Y)}+\frac{r}{X(X+rY)}),$$

$$L = 1-(\frac{\omega_{pe}^2}{\omega^2}\frac{\omega}{\omega+\Omega_e}+\frac{\omega_{pi}^2}{\omega^2}\frac{\omega}{\omega-\Omega_i}) = 1-(\frac{1}{X(X+Y)}+\frac{r}{X(X-rY)}),$$

so that, for completeness, we may write S and D as:

$$S = ½(R+L) = 1-(\frac{\omega_{pe}^2}{\omega^2-\Omega_e^2}+\frac{\omega_{pi}^2}{\omega^2-\Omega_i^2}) = 1-(\frac{1}{X^2-Y^2}+\frac{r}{X^2-r^2Y^2}),$$

and

$$D = ½(R-L) = -\frac{\omega_{pe}^2\Omega_e}{\omega(\omega^2-\Omega_e^2)}+\frac{\omega_{pi}^2\Omega_i}{\omega(\omega^2-\Omega_i^2)} = \frac{r^2Y}{X(X^2-r^2Y^2)}-\frac{Y}{X(X^2-Y^2)}.$$

The Appleton-Hartree equation is simply the dispersion relation, equation (5.6), in a different form and it may be constructed in the manner demonstrated below. If we add An^2 to both sides of equation (5.6), we get:

$$(An^2+(A-B))n^2 = An^2-C,$$

which may be rearranged to give an expression for the square of the refractive index:

$$n^2 = \frac{An^2-C}{An^2+A-B}. \tag{5.9}$$

Since equation (5.6) is a quadratic in n^2 it must have the usual solutions:

$$n^2 = \frac{B \pm \sqrt{B^2 - 4AC}}{2A},$$

which we now substitute into equation (5.9), yielding (after a little algebra):

$$n^2 = 1 - \frac{2(A - B + C)}{2A - B \pm \sqrt{B^2 - 4AC}}.$$

Finally, we replace A, B and C by the equivalent expressions for S, P, D, R and L in terms of the dimensionless variables X and Y, and define $\mu = \cos\theta$, $\nu = \sin\theta$ to obtain the Appleton-Hartree formulation of the cold plasma dispersion relation:

$$n^2 = 1 - \frac{2(X^2 - 1)}{2X^2(X^2 - 1) - \nu^2 X^2 Y^2 \pm (\nu^4 X^4 Y^4 + 4X^2 Y^2 (X^2 - 1)^2 \mu^2)^{1/2}}, \qquad (5.10)$$

where we have used the form of the dielectric tensor components which is appropriate for the case of frequencies sufficiently high to assume that the ions are stationary ($r = 0$).

There are certain similarities between the dispersion relation for the oscillations of the neutral atmosphere described by equation (4.12) and those of a cold plasma as described by equation (5.10). For the acoustic gravity waves, we define the analogous dimensionless variables $X = \omega/\omega_a$ and $Y = \omega_g/\omega_a$ so that we may rewrite equation (4.12) in an equivalent form to equation (5.10):

$$n^2 = \frac{X^2 - 1}{X^2 - \nu^2 Y^2}$$

where k represents the magnitude of the total wavevector and $\nu = \sin\phi$ with ϕ the angle which this vector makes with the vertical. In the special cases of vertical propagation or for $\omega \gg \omega_g$, the above equation describes a purely acoustic mode, independent of the gravitational field (when $X > 1$):

$$n^2 = 1 - \frac{1}{X^2}.$$

Similarly, for perpendicular propagation in a plasma or for $\omega \gg \Omega_e$, equation (5.10) reduces to the equation for the ordinary mode independent of the magnetic field:

$$n^2 = 1 - \frac{1}{X^2}.$$

6. The ordinary and extraordinary modes of perpendicular propagation

6.1. The ordinary mode

The mode described by $n^2 = P = 1 - \omega_{pe}^2/\omega^2$ is unaffected by the magnetic field and is therefore known as the *ordinary* mode. (This terminology has been adopted and modified from the field of crystal optics.) The dispersion relation for the ordinary mode is identical to that for transverse waves in an isotropic plasma. In terms of the new variables X and Y, we have:

$$P = 1 - \frac{1+r}{X^2}.$$

For small values of X, P is therefore very large and negative while as $X \to \infty$, $P \to 1$.

The ordinary mode undergoes a cutoff at $X = \sqrt{1+r}$, or $\omega^2 = \omega_{pe}^2(1+r)$, which is marginally greater than the electron plasma frequency (since the ratio of electron to ion masses is of the order of 5×10^{-4}). The variation of P as a function of frequency is shown schematically in Figure (5.2). Below the cutoff frequency, the square of the refractive index for this mode is always negative so that the corresponding wavenumber is pure imaginary and the wave is not a propagating solution. Regions in which a particular mode is allowed (i.e. where $n \in I\!R$) are called *pass bands* whereas regions where the mode is forbidden are its *stop* bands. The ordinary mode therefore has a stop band until just below the electron cyclotron frequency and a pass band beyond.

6.2. The extraordinary mode

The second perpendicular mode with $n^2 = RL/S$ is called the *extraordinary* mode. Although the R and L modes for parallel propagation have resonances at the electron cyclotron frequency, $X = Y$, and the ion cyclotron frequency, $X = rY$, respectively, the extraordinary mode is not resonant at either of these frequencies since S, being constructed from R and L, contains identical terms which cancel their effect. The extraordinary mode is then given by:

$$n^2 = \frac{(X^2+XY(r-1)-rY^2-1-r)(X^2-XY(r-1)-rY^2-1-r)}{(X^2-Y^2)(X^2-r^2Y^2)-(1+r)(X^2-rY^2)}. \qquad (5.11)$$

A typical plot of RL/S against frequency is illustrated in Figure (5.3), showing the cutoff and resonance frequencies calculated below. As $X \to 0$, the extraordinary mode tends to the limit given by $1 + (\omega_{pe}^2 + \omega_{pi}^2)/\Omega_e\Omega_i$, and at very high frequencies, $RL/S \to 1$. The cutoffs can easily be seen to be at the frequencies which satisfy:

$$X^2 \pm XY(r-1) - rY^2 - 1 - r = 0.$$

These frequencies are, in fact, the same as the cutoffs for the right and left circularly polarised for waves propagating parallel to the magnetic field - a fact which may easily be deduced by rearranging the expressions R, $L = 0$, using the definitions of the dielectric tensor elements given in §2. Because of their association with the *RCP* and *LCP* modes, these frequencies are denoted by ω_R for the upper sign and ω_L for the lower.

Since the factor r is three to four orders of magnitude smaller than 1, we will henceforth ignore all terms of order r, and higher, compared to unity. (Notice that this is *not* the same as assuming that the ions are stationary - we are still including ion effects.) Under this approximation, the equation giving the positions of the zeros of RL/S reduces to:

$$X^2 \mp XY - rY^2 - 1 = 0.$$

This quadratic is equivalent, using the initial dimensional variables, to:

$$\omega^2 \mp \omega\Omega_e - r\Omega_e^2 - \omega_{pe}^2 = 0 \quad \text{or} \quad \omega^2 \mp \omega\Omega_e - \Omega_e\Omega_i - \omega_{pe}^2 = 0,$$

where we have used the fact that the cyclotron and plasma frequencies of each species are related simply by the mass ratio i.e. $r\Omega_e = \Omega_i$.

Similarly, the resonances occur at the zeros of the denominator of equation (5.11) which, taking $r \ll 1$, are the frequencies satisfying:

$$X^4 - X^2(1+Y^2) + rY^2(1+rY^2) = 0.$$

The roots of this quadratic are:

$$X^2 = \tfrac{1}{2}(1+Y^2) \pm \tfrac{1}{2}(1+Y^2)\left(1 - \frac{4rY^2(1+rY^2)}{(1+Y^2)^2}\right)^{\frac{1}{2}},$$

$$= \tfrac{1}{2}(1+Y^2)\left(1 \pm \left(1 - \frac{2rY^2(1+rY^2)}{(1+Y^2)^2} \cdots\right)\right),$$

so that the extraordinary mode resonances occur at $X^2 \approx 1+Y^2$ and $X^2 \approx \dfrac{rY^2(1+rY^2)}{1+Y^2}$.

The first of these resonant frequencies is called the *upper hybrid* frequency and is usually written $\omega_{uh}^2 = \omega_{pe}^2 + \Omega_{pe}^2$. This frequency is slightly greater then the electron plasma frequency. For $n = 10^{19}\text{m}^{-3}$:

$$\frac{\Omega_e^2}{\omega_{pe}^2} = \frac{e^2B^2}{m_e^2}\frac{\varepsilon_0 m_e}{ne^2} \approx B^2,$$

so that, unless the magnetic field is exceptionally strong, the upper hybrid resonance will be very close to the plasma frequency. The other resonance, the *lower hybrid* frequency, may be expressed in terms of the plasma and gyro-frequencies as:

$$\omega_{lh}^2 = \frac{\Omega_e \Omega_i (\omega_{pe}^2 + \Omega_e \Omega_i)}{\omega_{pe}^2 + \Omega_e^2},$$

but this frequency is so low as to be often neglected.

An alternative form of equation (5.11) which emphasises the fact that the refractive index passes through two cutoffs and two resonances, allowing for a greater number of pass and stop bands than was present in the case of the ordinary mode, is:

$$n^2 = \frac{(\omega^2 - \omega_R^2)(\omega^2 - \omega_L^2)}{(\omega^2 - \omega_{lh}^2)(\omega^2 - \omega_{uh}^2)}. \tag{5.12}$$

If an even number out of the four terms on the right hand side of equation (5.12) is positive then n^2 will be positive and the extraordinary mode will propagate. Alternatively, stop bands will occur when an odd number of the conditions in equation (5.12) is positive. The relative positions of the cutoffs and resonances for this mode alter depending on the applied magnetic field strength.

7. Two special cold plasma waves

The dispersion relation, equation (5.6), is often simplified by the use of appropriate approximations in order to examine the propagation characteristics of special classes of waves. The result is a classification and nomenclature which is a subdivision of the fast and slow modes but which are only valid in certain limited regions of the CMA diagram. Many of these wave phenomena were observed experimentally before a fully comprehensive theoretical explanation of plasma waves was developed. We will only consider two examples for illustration - *Alfvén waves* and *Whistlers*, the former because of the overlap which they provide with the predominant wave motions of MHD, and the latter because of the similarity to the gravity waves of Chapter 4 which they exhibit.

7.1. Alfvén waves

These waves are a low frequency phenomenon, occurring predominantly below the ion cyclotron frequency, and are particularly important in MHD where their characteristic velocity plays a central role in all three wave modes - the shear Alfvén and fast and slow magnetosonic waves. We derive the dispersion relation for Alfvén waves from equation (5.6) under the assumption that $\omega \ll \Omega_i$. Applying this assumption to the dielectric tensor elements enables us to approximate their form and hence establish the equation for the refractive index which describes the wave motion. In terms of the dimensionless variables $X = \omega/\omega_{pe}$ and $Y = \Omega_e/\omega_{pe}$ introduced earlier, we have:

$$R,L = 1-\left(\frac{1}{X(X\mp Y)}+\frac{r}{X(X\pm rY)}\right),$$

which may be expanded for $X \ll rY$ to give:

$$R,L \approx 1+\frac{(1+r)}{rY^2}\left(1\mp\frac{X}{rY}\right)\left(1\pm\frac{X}{Y}\right),$$

$$\approx 1+\frac{(1+r)}{rY^2},$$

so that $S \approx R \approx L$ and $D \approx 0$. In this notation, P was written $P = 1-\frac{1+r}{X^2}$. Taking the ratio of the second terms in the expressions for P and R gives:

$$\frac{1+r}{X^2}\frac{rY^2}{1+r} = \frac{rY^2}{X^2} = \frac{\Omega_i^2}{r^2\omega^2} \gg 1,$$

with the result that $|P| \gg |R|,|L|,|S|$ and so, except in the neighbourhood of perpendicular propagation (in the notation of equation (5.6)):

$$A \approx -\frac{1+r}{X^2}\cos^2\theta,$$

$$B \approx -\frac{1+r}{X^2}(1+\cos^2\theta)\left(1+\frac{1+r}{rY^2}\right),$$

$$C \approx -\frac{1+r}{X^2}\left(1+\frac{1+r}{rY^2}\right)^2.$$

By equation (5.6), the refractive indices must therefore be given by:

$$n^2 \approx 1+\frac{(1+r)}{rY^2} \quad \text{and} \quad n^2 \approx \frac{1}{\cos^2\theta}\left(1+\frac{1+r}{rY^2}\right).$$

Noticing that $(1+r)/rY^2 = (\omega_{pe}^2+\omega_p^2)/\Omega_e\Omega_i$, and using the overall charge neutrality of the plasma, $n_e = n_i$, we may rewrite the equations for the refractive indices:

$$n^2 = 1+\frac{\rho}{\varepsilon_0 B^2} \quad \text{and} \quad n^2 = \frac{1}{\cos^2\theta}\left(1+\frac{\rho}{\varepsilon_0 B^2}\right), \tag{5.13}$$

where $\rho = n_i m_i + n_e m_e$ is the mass density of the plasma defined in Chapter 3. The important feature of these waves is their common phase velocity:

$$v_{ph} = \frac{c}{(1+\rho/\varepsilon_0 B^2)^{1/2}} = \frac{c\varepsilon_0^{1/2}B}{\rho^{1/2}(1+\varepsilon_0 B^2/\rho)^{1/2}} \approx \sqrt{\frac{B^2}{\mu_0\rho}},$$

which is called the *Alfvén velocity.* This may be recognised as the velocity of the waves postulated in Chapter 4 by considering the magnetic field lines to be analogous to strings under tension. We have thus introduced a degree of mathematical rigour into the description of the waves in equation (5.13) - the fast and slow, or compressional and shear, Alfvén waves respectively. These waves show how the results of MHD and cold plasma theory merge in the low pressure limit.

Because of their extremely low frequency, the Alfvén waves appear in region [12] of the CMA diagram (Figure (5.1)). The shear Alfven wave is drastically altered when it passes through the ion cyclotron resonance and emerges as an ion cyclotron wave. On the other hand, the compressional Alfvén wave suffers no radical change and continues to propagate for frequencies above this resonance, finally disappearing at the electron cyclotron resonance.

7.2. Whistlers

These ionospheric disturbances excited by lightning flashes are guided along the Earth's magnetic field lines to distant points on the Earth's surface where they may be detected. They were first observed early this century on sonograms which show the variation of frequency spectrum with arrival time and thus provide a diagnostic tool for analysis of ionospheric conditions. Because the pulses which are produced are rich in very low frequency components (10Hz to 100Hz), Whistler frequencies are often audible to the human ear as gliding whistles descending in pitch. In the Appleton-Hartree dispersion equation, which is appropriate to this frequency range, we take $\Omega_i \ll \omega < \Omega_e$ together with $\omega|\Omega_e| < \omega_{pe}^2$ and $\theta \sim 0°$, since propagation is nearly along the magnetic field lines, so that the ratio of the magnitudes of the two terms under the square root in equation (5.10) is:

$$\frac{2XY(X^2-1)\mu}{v^2X^2Y^2} = \left(\frac{X}{Y} - \frac{1}{XY}\right)\frac{2\mu}{v^2} \gg 1$$

and thus the lower sign of the dispersion relation, equation (5.10) reduces to:

$$n^2 = 1 - \frac{1}{X^2(1 - Y/X\mu)}. \quad (5.14)$$

Under the additional constraint, $\omega \ll \Omega_e \cos\theta$, the dispersion relation is further reduced to the form:

$$n^2 = 1 + \frac{1}{XY\mu} \approx \frac{1}{XY\mu}.$$

--

In other words, $\frac{ck}{\omega} = \frac{\omega_{pe}}{\sqrt{\omega\Omega_e\mu}}$. From this, we may calculate the group velocity to be:

$$v_g = \frac{\partial\omega}{\partial k} = \frac{2c}{\omega_{pe}}(\omega\Omega_e\cos\theta)^{1/2},$$

which demonstrates that higher frequencies will be detected by the receiver slightly before the lower ones. Whistlers are the manifestation of the *RCP* mode in the range of frequencies prescribed above and occur in region [8] of the CMA diagram. It is interesting to compare the Whistler wave dispersion relation in the form of equation (5.14) with the gravity wave formula:

$$n^2 = \frac{X^2-1}{X^2-v^2Y^2}.$$

In particular, these two waves have critical angles of propagation: $\sin\phi = \omega/\omega_g$ for the gravity wave and $\cos\theta = \omega/\Omega_e$ for the Whistler. The two main differences between these waves are that the energy propagation of Whistlers tends to be aligned along the magnetic field direction whereas energy propagation for gravity waves is confined to horizontal directions, perpendicular to the axis of symmetry. Secondly, although both waves are clearly anisotropic, gravity waves are not gyrotropic - they do not split into differently polarised components with different phase speeds.

8. Summary

Viewing the cold plasma as a dielectric medium containing currents, we have derived the general dispersion formula for cold plasma waves in several different forms, each of which has proved useful under certain circumstances. We have considered the cases of waves propagating parallel and perpendicular to the background magnetic field in detail and have introduced a non-dimensional notation which will be used extensively in the following chapter. The approximations necessary for the consideration of Alfvén waves, the Appleton-Hartree magnetoionic formula and Whistlers have been discussed in order to emphasise the possible widespread applications of cold plasma theory. Some similarities between certain cold plasma waves and those which propagate in the non-conducting medium of Chapter 4 are noted. We have now laid the foundation for considering waves in an inhomogeneous cold plasma.

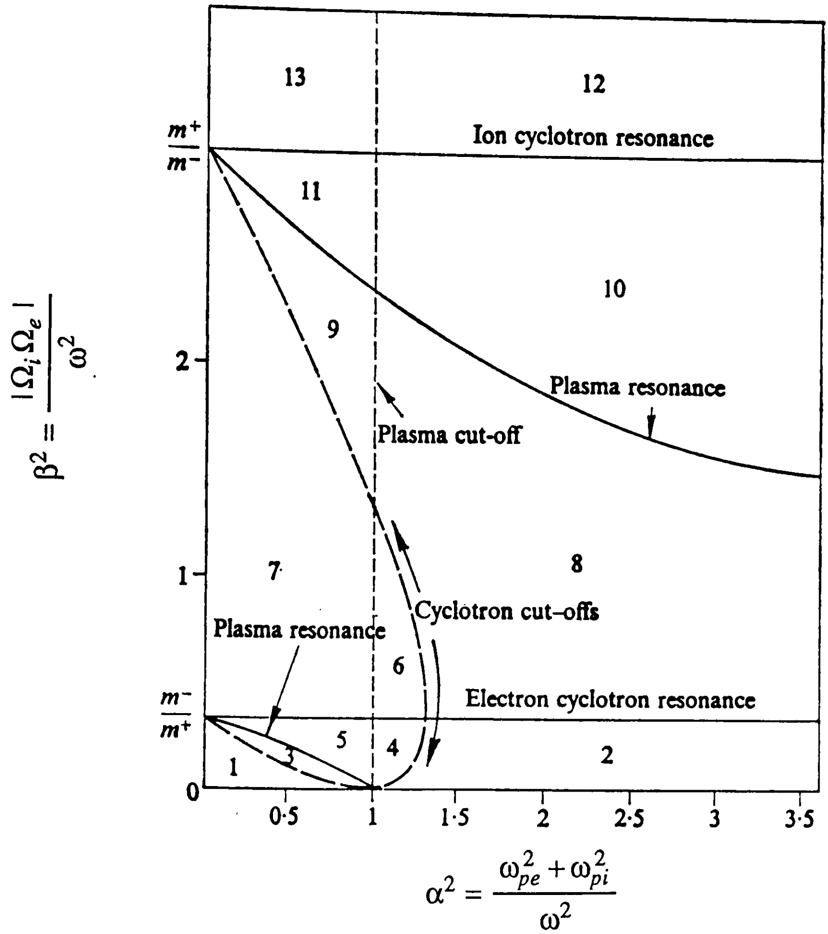


Figure (5.1) The CMA diagram showing the subdivision of parameter space by the ion cyclotron resonance ($L = \infty$), plasma resonance ($S = 0$), electron cyclotron resonance ($R = \infty$), plasma cutoff ($P = 0$) and cyclotron cutoff curves (R, $L = 0$) into thirteen regions in each of which the wave-normal surfaces are topologically distinct. Regions whose characteristic numbers differ by 6 have similar properties since the electron behaviour at high frequencies is mirrored by the ions at low frequencies. (The ratio m_i/m_e shown is unrealistically low for a gaseous plasma but not necessarily for a solid state plasma.)

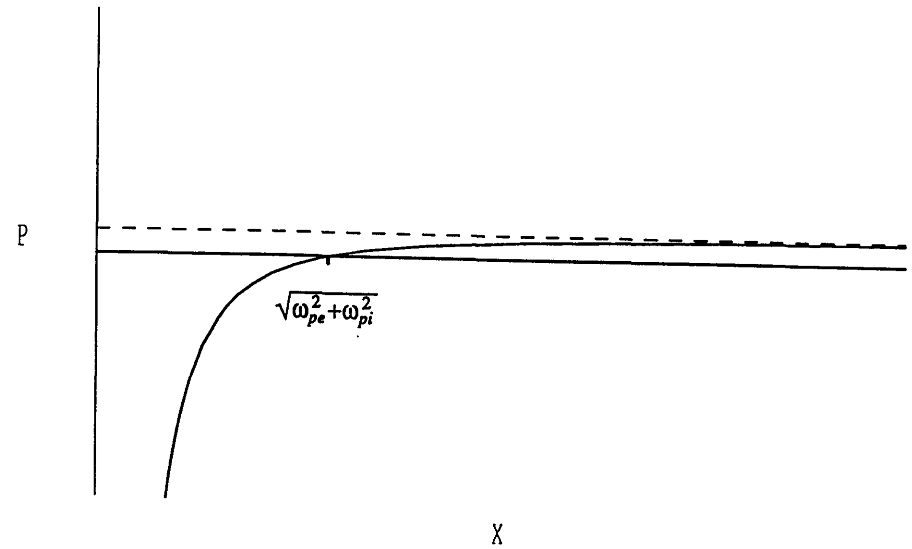


Figure (5.2) Schematic diagram of the variation of the ordinary mode of perpendicular propagation in a cold plasma, P, with dimensionless frequency, $X = \frac{\omega}{\omega_{pe}}$.

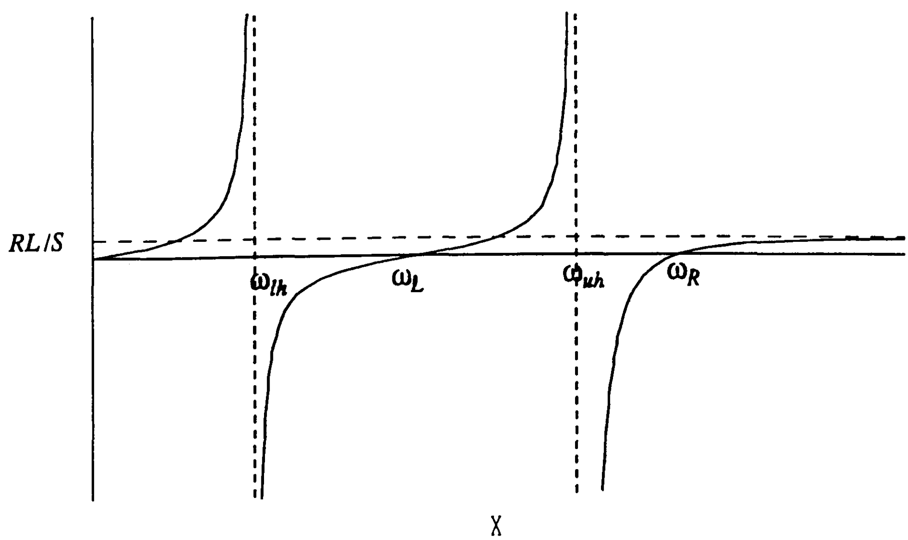


Figure (5.3) Schematic diagram of the variation of the extraordinary mode, RL/S, with frequency.

Chapter 6 - Propagation of cold plasma waves in a spatially-rotating magnetic field

1. Introduction

Real plasmas seldom closely resemble the highly-idealised cold plasma model used in Chapter 5, but there are strong arguments in favour of continuing and extending it. This representation yields a good, basic description of the plasma relatively quickly and without resort to extremely complicated mathematics. Thus we can construct a physical picture upon which to build when finite-temperature effects are included. An example of such an extension was given in the previous chapter in connection with plasma oscillations along the magnetic field lines. Although solutions of the dispersion relation, equation (5.7), these were seen merely to be internal oscillations of the cold plasma, only becoming truly wavelike solutions when thermal effects were taken into consideration.

Another good reason for studying the cold plasma model is that it enables us to determine solutions analytically. These are extremely important as a comparison for computational results which are restricted to one set of parameters at a time and which therefore provide less of a global view. In addition, physically interesting results may be lost or overlooked in performing complicated numerical calculations. We feel justified in using the cold plasma model because of the wealth of valuable information it has provided in the past, despite its apparently restrictive assumption of zero thermal velocity, and because of its appealing simplicity.

In this chapter we will extend the model discussed in Chapter 5 to study a cold plasma whose properties are not constant but are allowed to vary with position. The study of spatially varying plasmas is one way in which we may increase the realism of our model since plasmas, whether man-made or natural, will be inhomogeneous to some extent, as was discussed in Chapter 1. There, particular attention was paid to plasmas with varying magnetic fields but the possibility of gradients in other plasma properties, e.g. density and temperature, was also noted.

We choose to make our plasma inhomogeneous by the introduction of a non-uniform magnetic field which replaces the constant field of Chapter 5 and then observe the consequent changes in the behaviour of waves in the plasma. It will be shown that, as a result of this primary variation, even the simplest possible equilibrium configuration is vastly more complicated than the one obtained previously, containing a non-zero velocity (and hence current) arising naturally from our initial conditions. The plasma can no longer remain stationary initially but must flow with a velocity which is dependent on the magnitude of the magnetic field and its degree of rotation. We will show that this variation introduces completely new phenomena into our description of the waves in the cold plasma and that, although superficially very similar, the model in this chapter differs markedly from that of Chapter 5.

2. The model equations

The model equations describing our system are those appropriate to a two-species cold plasma which were used in the previous chapter but, to allow for new effects introduced by the inhomogeneity, we must return to an earlier, more general statement of them. It will prove necessary in subsequent sections to retain the continuity equation in addition to those equations already required in our study of the homogeneous case and we therefore apply the cold plasma approximation directly to equations (3.4) and (3.5), restating the results in full here along with Maxwell's equations (3.8) and (3.9). In this chapter, we again assume that the plasma is collisionless - there are no processes of particle creation or destruction, $S_s = 0$ - and that the contributions to the momentum equation due to collisions, $\mathbf{A}_s$ is also set equal to zero, which is equivalent to taking the collision frequency, $\nu = 0$. Our complete set of (unlinearised) model equations is then:

$$\frac{Dn_s}{Dt} = -n_s \nabla\cdot\mathbf{v}_s , \tag{6.1}$$

$$m_s \frac{D\mathbf{v}_s}{Dt} = q_s(\mathbf{E}+\mathbf{v}_s\times\mathbf{B}) , \tag{6.2}$$

$$\nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t} , \tag{6.3}$$

$$\nabla\times\mathbf{B} = \mu_0 (\mathbf{J}+\varepsilon_0\frac{\partial\mathbf{E}}{\partial t}) , \tag{6.4}$$

$$\mathbf{J} = \sum_s n_s\, q_s\, \mathbf{v}_s , \tag{6.5}$$

where the variables have been defined previously and we have expressed equations (6.1) and (6.2) in terms of the advective derivative which we discussed in Chapter 3:

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v}_s\cdot\nabla .$$

From this basic description, we will construct a wave equation following the pattern of Chapter 5. First, we will decide upon the equilibrium state of the plasma and then we will proceed to linearise equations (6.1) to (6.5) in the usual way. Following this, the perturbed variables will be eliminated in turn until only the components of the first order electric field remain.

3. The equilibrium configuration

A clear understanding of the geometry of the problem which we are considering is crucial in order to follow the physical processes involved. The precise alignment and variation of the background magnetic field determine the orientation and motion (if any) of the plasma constituents. On the other hand, the equilibrium magnetic field configuration is to a large extent imposed upon us by our desire to construct an equilibrium which satisfies equations (6.1) to (6.5).

The case of plasma wave propagation at a fixed angle to a magnetic field of varying strength - the magnetic beach model - has been considered by other authors (e.g. Diver, 1986). A logical progression is therefore to consider a magnetic field with changing orientation so that the wave continually "sees" the field in a different direction. We will therefore be concerned with a wave propagating perpendicularly to a magnetic field which changes orientation along the direction of propagation (see Figure (6.1)).

Fidone and Granata (1971) consider the observed excitation of a small ordinary wave component by an extraordinary wave propagating across the magnetic field of a toroidal configuration, as it passes through the upper hybrid resonance. By assuming that the magnetic field lies in the plane perpendicular to the direction of propagation, they are in essence supposing that the poloidal component of the total magnetic field is sufficiently small as to be neglected, but that it contributes an overall twisting to the field lines. Although the geometry of their problem is therefore very similar to ours, the results of their method apply, at best, only to a slowly varying magnetic field whereas we consider an arbitrarily large magnetic field gradient. The main drawbacks with Fidone and Granata's paper - the use of Budden's equations throughout and the neglect of the equilibrium current - were summarised in Chapter 2. The need for an equilibrium current to balance the changing magnetic field will be amply demonstrated later in this section.

More recently, Choudhury (1988) has extended the work of Grossman and Weitzner (1984) to examine the possibility of heating tokamaks at frequencies near the lower hybrid resonance. Both works attempt to explain why the core region of experimental plasmas appears to be accessible to lower hybrid waves, contrary to the predictions of cold plasma theory and geometrical optics. Grossman and Weitzner found that inclusion of a density gradient increased the accessibility to the plasma interior of these waves when the modes propagated perpendicular to the ambient magnetic field. Choudhury includes the effects of a varying magnetic field (and thus an equilibrium plasma current) and parallel propagation to generalise these results. We will deal with individual points of similarity between our work and that of Choudhury as they arise during the course of analysis. The main difference between the approaches is that, whereas we select a specific magnetic field geometry at the outset, Choudhury attempts to cover a broader class of field configurations which results in a trade-off between choosing a general initial description and being forced to make assumptions at later stages. Although we will be restricted to definite variation of the equilibrium variables, we will succeed in deriving an analytic solution for our model. Choudhury, on the contrary, must continually make approximations in order to progress and finally derives a dispersion relation which is solved numerically for the case of a homogeneous plasma.

The geometry in which we are interested bears a closer resemblance to that of the reverse field pinch discussed in Chapter 1 than to a tokamak. Near the point of field reversal, the toroidal field is very small while there remains a significant poloidal component which is constantly changing orientation with respect to the minor axis of the torus. Although we assume initially that the magnitude of the field varies as well as its direction, we will proceed to exclude this possibility from our analysis using the periodic form of the field's direction cosines. Intuitively, the work of Chapter 5 would then suggest that the allowed modes of propagation should be related to the modes permissible for parallel and perpendicular propagation in a homogeneous plasma - the ordinary and extraordinary modes.

As before, we shall seek the fundamental modes of an unlimited plasma. In the context of solving a particular problem under a set of well-defined, experimentally-derived, external conditions, the plasma waves would be further constrained to a discrete subset of the continuous set of allowed oscillations derived in the remainder of this chapter. Since we will only be interested in solving our problem in the context of an infinite plasma, our rotating field must obviously repeat its configuration at set distances throughout the system. We therefore choose the simplest description of a 2π–periodic rotation available and express our ambient magnetic field in terms of trigonometric functions of position (cos and sin). From the initial requirement that the field should vary periodically in space, we will show that this variation must conform to certain strict rules.

We are interested in a static (but not necessarily stationary) equilibrium, so that none of the zero-order quantities are time-dependent. Thus, the right hand side of equation (6.3), together with the displacement current term in equation (6.4), vanish in equilibrium. We recognise that the curl operator acting on the equilibrium magnetic field is non-zero now that the field is not spatially uniform and, in order to balance equation (6.4), we must therefore include a non-vanishing right hand side. This non-zero $\mathbf{J}_0$ represents an equilibrium current, the need for which in inhomogeneous plasmas has been recognised by a number of authors in this field, including Choudhury, and Lashmore-Davies and Stenflo (1981) in a paper on the MHD stability of a helical magnetic field of arbitrary amplitude. (Here, helical denotes $\mathbf{B}_0 = B_{z0}\,\hat{\mathbf{z}} + B_{\perp 0}\,[-\hat{\mathbf{x}}\sin(\omega_0 t - k_0 z) + \hat{\mathbf{y}}\cos(\omega_0 t - k_0 z)]$), where B_{z0}, $B_{\perp 0}$, ω_0 and k_0 are constants.

The equilibrium current arises from the flow of charged particles and so we must also introduce a corresponding non-zero equilibrium velocity into our system, as implied by equation (6.5). We combine equations (6.1), (6.2), (6.4) and (6.5) below to show that the current is constrained to flow in one of two opposite directions. (We continue to suppose that the ions are singly charged for simplicity so that $n_e = n_i = n_0$ in equilibrium.)

In equilibrium, equation (6.1) becomes:

$$\mathbf{v}_0 \cdot \nabla n_0 = -n_0 \nabla \cdot \mathbf{v}_0,$$

the left hand side of which is zero since we assume that the ambient number density of each species is fixed. Choosing the vertical axis to be the only direction along which there is any variation, so that z is our only independent variable, the resulting restriction on the velocity may be written $\partial v_{0sz}/\partial z = 0$. (The arguments for and against a single independent variable will be presented in the next section.) This constraint on the z-component of the velocity also means that the total derivative in equation (6.2) is identically zero in equilibrium. Writing equation (6.2) explicitly for both species:

$$\mathbf{v}_{0i} \times \mathbf{B}_0 + \mathbf{E}_0 = 0, \tag{6.2a}$$

$$\mathbf{v}_{0e} \times \mathbf{B}_0 + \mathbf{E}_0 = 0, \tag{6.2b}$$

and subtracting equation (6.2b) from (6.2a), we obtain:

$$(\mathbf{v}_{0i} - \mathbf{v}_{0e}) \times \mathbf{B}_0 = 0.$$

In terms of the common background number density and the electronic charge, equation (6.5) may be written as:

$$\mathbf{J}_0 = n_0 e (\mathbf{v}_{0i} - \mathbf{v}_{0e}),$$

so that the equilibrium current must be parallel to the magnetic field. The current and the magnetic field are also related via equation (6.4) and, in order to satisfy both these relations simultaneously, the magnetic field must be of the form:

$$\nabla \times \mathbf{B}_0 = \lambda \mathbf{B}_0, \tag{6.6}$$

which is the equation of a *force-free* magnetic field. Equation (6.6) implies that, as we travel along a field line in the direction of $\mathbf{B}_0$, the neighbouring field lines curl in a fixed sense round that line and so a force-free field is essentially a twisted field. Such fields are especially important in *low*-β plasmas in which the gas pressure is small compared with the magnetic pressure. They are appropriate to certain types of toroidal fusion experiments, where the equilibrium surfaces formed are nested tori, and often occur in astrophysical plasmas. Chandrasekhar and Woltjer (1958) explained the extremely regular magnetic field of the Crab Nebula in terms of a force-free field. They argued that, because the interstellar medium is almost perfectly conducting, any currents present would give rise to a $\mathbf{J}\times\mathbf{B}$ force of such magnitude that it could not be balanced by gravitational or pressure forces. The only way to construct an equilibrium configuration would therefore be by aligning the current along the magnetic field direction, resulting in a force-free situation. They also concluded from certain thermodynamic arguments that force-free magnetic fields are highly likely to occur in interstellar space. Nakagawa and Raadu (1972) used the likelihood of naturally occurring force-free fields in their discussion of the magnetic field of the Sun's chromosphere and low corona. By employing illustrations of the striking similarity between solar H_α observations and the lines of force of force-free magnetic fields, they built a plausible case for the magnetic field in these regions being force-free. From this assumption, they then demonstrated a comparatively simple method of determining the magnetic field uniquely from its observed vertical component in the case of λ = *constant*.

In proving a variational principle about force-free fields, namely that force-free fields with constant λ represent the state of lowest magnetic energy in a closed system, Woltjer (1958) paved the way for subsequent investigations. Workers in the fusion field were at first surprised by the spontaneous generation of reverse fields in ZETA and other pinch machines and the stability of the resulting configurations. These anomalies were explained in a classic paper by Taylor (1974) where he demonstrated that both phenomena were related to the relaxation, under constraint, of a magnetic field to a final force-free state. He showed that a perfectly conducting plasma, for which all topological properties of the field lines must remain invariant, would relax to one of an infinite number of equivalent minimum energy states - force- free states characterised by different values of λ. (Note that by taking the divergence of equation (6.6), it can easily be shown that λ is either zero or constant along a field line. It need not, however, be constant for different field lines.) Taylor then relaxed the assumption of perfect conductivity and showed that the final minimum energy state was in this case unique, with λ the same on all field lines. Another interesting result about force-free fields was derived in an elegant paper by Jette (1970). He proved that, in resistive MHD, the only force-free magnetic fields which remain force-free in time are those for which λ is constant in both space and time.

3.1. Restrictions imposed by the force-free condition of the ambient magnetic field

From our original loosely defined prescription of a varying magnetic field, our model has dictated that this field must be force-free. We now use the periodicity of the field, alluded to above, to complete its specification. If the magnetic field varies periodically along one axis of a plane, it must also vary in a compensatory fashion along the other. With the single independent variable z, we write $B_{0x} = B_0(z)\cos\phi(z)$ and $B_{0y} = B_0(z)\sin\phi(z)$ where $B_0(z)$ is the magnitude of the magnetic field. Suppose our field has an arbitrary third component, e.g.

$$\mathbf{B}_0 = (B_0\mu, B_0\nu, b), \tag{6.7}$$

where we have introduced the notation of direction cosines, $\mu = \cos\phi(z)$ and $\nu = \sin\phi(z)$ which we will employ henceforth. Clearly, the third component of $\mathbf{B}_0$, b, must be a constant in order to satisfy $\nabla\cdot\mathbf{B}_0 = 0$. We now use equation (6.4) to show that b must in fact vanish. For the left hand side of equation (6.4), we have:

$$\nabla\times\mathbf{B}_0 = \det\begin{bmatrix} \hat{\underline{x}} & \hat{\underline{y}} & \hat{\underline{z}} \\ 0 & 0 & \dfrac{\partial}{\partial z} \\ B_0\mu & B_0\nu & b \end{bmatrix},$$

$$= (-(B_0\nu)', (B_0\mu)', 0),$$

$$= B_0'(-\nu,\mu,0) - B_0\phi'(\mu,\nu,0),$$

where $' \stackrel{\partial}{=} \partial/\partial z$.

To ensure that the magnetic field satisfies the force-free condition as required by equation (6.6), (that is, to make this vector parallel over all space to the magnetic field vector given by equation (6.7)), the magnitude of the ambient field must be constant and its z-component must be zero i.e. $B_0' = b = 0$. We may now identify λ in equation (6.6) with $-\partial\phi/\partial z$ and note that this need not be constant throughout the plasma although it is obviously constant along each field line.

These restrictions are placed on b and B_0 by our use of a single independent variable and could be made less stringent by allowing variation in at least one other direction. We would then be faced with a set of coupled, partial differential equations to solve. The complications resulting in other parts of the analysis and the increased difficulty incurred in visualising the more complex geometry do not recommend this step, at least not until the present approach has been fully explored.

In summary, the magnetic field no longer points in a fixed direction as it did in Chapter 5. For any position ($\mathbf{z}$), the magnetic field direction is specified by the angle it makes with the x-axis (ϕ) and this angle changes as we go along $\mathbf{z}$. We have also shown that, having selected a magnetic field with periodic variation in the plane perpendicular to $\mathbf{z}$, the only possible equilibrium permitted by the equations of the cold plasma model is one in which the remaining component of the field is zero and the magnitude of the field is constant. As a result of the above arguments, the magnetic field will henceforth be taken to be:

$$\mathbf{B}_0 = B_0(\mu(z),\nu(z),0),$$

where B_0 is constant.

3.2. Equilibrium velocities in terms of $\mathbf{B}_0$

The general form of the relation between $\mathbf{B}_0$ and $\mathbf{v}_{0s}$ has already been used to demonstrate the force-free nature of the magnetic field. Their exact interdependence will now be examined and a linear expression for $\mathbf{v}_{0s}$ in terms of $\mathbf{B}_0$ derived. It was shown above that

$$(\mathbf{v}_{0i} - \mathbf{v}_{0e}) \times \mathbf{B}_0 = 0.$$

There are two possible ways of satisfying this relation, either:-

(i) both $\mathbf{v}_{0e}$ and $\mathbf{v}_{0i}$ are individually parallel to $\mathbf{B}_0$, or

(ii) the difference of the two vector velocities $(\mathbf{v}_{0i} - \mathbf{v}_{0e})$ is parallel to the field without either being in that direction.

In the latter case, equation (6.2) requires that an external electric field be applied, satisfying

$$\mathbf{E}_0 = -\mathbf{v}_{0i} \times \mathbf{B}_0 = -\mathbf{v}_{0e} \times \mathbf{B}_0.$$

We now show that it is not possible to construct an equilibrium containing such a "Lorentz" electric field under the assumptions of our model. With a view to a contradiction, we suppose there exists a non-zero electric field satisfying equations (6.2a) and (6.2b) simultaneously, then by equation (6.3) it must also satisfy $\nabla\times\mathbf{E}_0 = -\nabla\times(\mathbf{v}_{0s}\times\mathbf{B}_0) = 0$. Applying the well-known identity theorem relating vector and scalar products:

$$\nabla \times (\mathbf{A}\times\mathbf{B}) = (\mathbf{B}\cdot\nabla)\mathbf{A} - (\mathbf{A}\cdot\nabla)\mathbf{B} + (\nabla\cdot\mathbf{B})\mathbf{A} - (\nabla\cdot\mathbf{A})\mathbf{B},$$

to the equilibrium form of equation (6.3) yields

$$(\mathbf{B}_0\cdot\nabla)\mathbf{v}_{0s}-(\mathbf{v}_{0s}\cdot\nabla)\mathbf{B}_0+(\nabla\cdot\mathbf{B}_0)\mathbf{v}_{0s}-(\nabla\cdot\mathbf{v}_{0s})\mathbf{B}_0=0.$$

Now, $\nabla\cdot\mathbf{B}_0=0$ and the magnetic field rotates in a plane perpendicular to the direction of variation so that we also have $(\mathbf{B}_0\cdot\nabla)=0$, leaving

$$(\mathbf{v}_{0s}\cdot\nabla)\mathbf{B}_0+(\nabla\cdot\mathbf{v}_{0s})\mathbf{B}_0=0.$$

Since the z-component of the magnetic field is zero, this vector equation contains only two non-zero parts:

$$v_{0sz}\frac{\partial B_{0x}}{\partial z}+\frac{\partial v_{0sz}}{\partial z}B_{0x}=0,$$

and

$$v_{0sz}\frac{\partial B_{0y}}{\partial z}+\frac{\partial v_{0sz}}{\partial z}B_{0y}=0.$$

On inserting the known form of the components of $\mathbf{B}_0$, we obtain from these respectively

$$-v_{0sz}B_0\phi'\sin\phi+\frac{\partial v_{0sz}}{\partial z}B_0\cos\phi=0,$$

and

$$v_{0sz}B_0\phi'\cos\phi+\frac{\partial v_{0sz}}{\partial z}B_0\sin\phi=0,$$

which on rearranging become:

$$\frac{1}{v_{0sz}}\frac{\partial v_{0sz}}{\partial z}=\tan\phi\frac{\partial\phi}{\partial z},$$

and

$$\frac{1}{v_{0sz}}\frac{\partial v_{0sz}}{\partial z}=-\frac{1}{\tan\phi}\frac{\partial\phi}{\partial z}.$$

Equating the right hand sides of these two equations results in:

$$\tan^2\phi=-1 \quad \text{or} \quad \sin^2\phi+\cos^2\phi=0,$$

which is plainly impossible and so the original assumption of a non-trivial electric field has been contradicted. Thus we must adhere to the conditions of option (i) where $\mathbf{E}_0=0$ and the flow velocities of both species are parallel to $\mathbf{B}_0$. We write:

$$\mathbf{v}_{0i} = \alpha(\mu(z), \nu(z), 0) \text{ and } \mathbf{v}_{0e} = \beta(\mu(z), \nu(z), 0),$$

where $\alpha - \beta = -(\frac{\partial \phi}{\partial z} B_0)/(\mu_0 n_0 e)$ on substituting in (6.4) from (6.5). Whenever the magnetic field is sheared and non-vanishing (i.e $\partial\phi/\partial z \neq 0$ and $B_0 \neq 0$) this expression shows that α must differ from β.

There thus remains one parameter which we are free to choose before the equilibrium velocities are completely specified in terms of the magnetic field. We set α (or β) by invoking another physical argument, namely that the nett momentum of the system as a whole be zero so that $\mathbf{v}_{0i} = -(m_e \mathbf{v}_{0e})/m_i$. The parameters introduced above then become

$$\alpha = \frac{-\frac{\partial \phi}{\partial z} B_0 m_e}{\mu_0 n_0 e (m_e + m_i)} \text{ and } \beta = \frac{\frac{\partial \phi}{\partial z} B_0 m_i}{\mu_0 n_0 e (m_e + m_i)}.$$

Attention is often restricted to electron motions because the ions are so much more massive - the ions' inertia inhibits their response to all but the lowest frequencies (e.g. Budden (1961) and Fidone and Granata (1971)). This is equivalent to taking $\alpha = 0$ and $\beta = \frac{\partial \phi}{\partial z} B_0/(\mu_0 n_0 e)$. We do not employ this approximation since solutions derived without it will be applicable to a wider frequency range (including $\omega < \Omega_i$, the ion cyclotron frequency, below which the approximation of fixed ions is no longer appropriate) and our use of computer algebra (REDUCE (1987)) eases much of the algebraic burden entailed in retaining both species. To recap, we now have:

$$\mathbf{v}_{0e} = \frac{\mathbf{B}_0}{\mu_0 n e (1 + m_e/m_i)} \frac{\partial \phi}{\partial z},$$

$$\mathbf{v}_{0i} = -\frac{m_e}{m_i} \mathbf{v}_{0e},$$

which may combined into $\mathbf{v}_{0s} = A_s \mathbf{B}_0 \partial\phi/\partial z$, where the A_s are species-dependent constants defined by inspection of the above.

If we now compare the results from this section with those of Choudhury's paper , we see that although the two systems share many of their properties, our more stringent definition of the magnetic field has led to a unique statement of the equilibrium configuration while Choudhury has been forced to make further choices about the remaining equilibrium variables. (Our problem serves as a reminder that the compatibility of equilibrium variables must always be cross-checked in individual cases.) We have been compelled by our original choice of field variation to consider a force-free magnetic field whereas Choudhury includes a force-free field, since it is a plausible geometry for the low-beta regime of operation in which he is interested. Similarly, he assumes that $\mathbf{E}_0 = 0$, but we have demonstrated that we must have this in order to balance $\mathbf{v}\times\mathbf{B}_0$ for both species. Finally, he supposes that $\mathbf{v}_{0i} = 0$ so that there is no mass flow in equilibrium and the only contribution to the current comes from the electrons. This contrasts with our requirement that the total momentum of the system must be conserved - an assumption which seems to be more easily justified. The fact that the equilibrium current and electron velocity are parallel to the magnetic field is, however, common to both descriptions.

4. The perturbed system

We aim to construct and solve a wave equation in the electric field components as in the last chapter and we develop a similar framework by considering the plasma to be a dielectric material. It was demonstrated in the earlier sections of this chapter that there are now fewer equilibrium quantities which are identically zero and the perturbation expansions of our model equations will therefore contain more non-vanishing terms. As a result, it will not prove possible to write the plasma current as a product of a conductivity tensor and the perturbed electric field alone - instead, the first-order current will become a sum of three terms involving the perturbed electric field plus its first two derivatives.

Having related all the equilibrium quantities, apart from the density, to the prescribed magnetic field, we now substitute these expressions in the perturbed forms of equations (6.1) to (6.5) and linearise in the usual way. One consequence of the non-zero velocities in equilibrium which we must now consider is the possible inclusion of density changes in our description of the plasma. The linearised form of equation (6.5) contains a term $n_1\mathbf{v}_{0s}$ which was absent in the uniform ($\mathbf{v}_0 = 0$) case. Since the homogeneous model involved only the background density common to both ions and electrons, this is our first encounter with density variation and we must therefore include an extra evolution equation to account for such density fluctuations and consequently must retain equation (6.1) in our description of the system. The linearised equations are derived in the succeeding sections.

4.1. Continuity equation

The continuity equation describing the evolution of the number density of each species is essential to close the linearised set of equations appropriate to a cold plasma containing a rotating magnetic field. Since our aim is again to eliminate all variables in favour of $\mathbf{E}$, we express the perturbed number density in terms of the velocity, which may in turn be related to the electric field via the momentum equation, (6.2).

Rewriting the total derivative in the equation of continuity, equation (6.1), as the sum of its two constituent parts gives:

$$\frac{\partial n_s}{\partial t} + \mathbf{v}_s \cdot \nabla n_s + n_s \nabla \cdot \mathbf{v}_s = 0,$$

and taking each variable to be composed of equilibrium and perturbed components, we have:

$$\frac{\partial n_{1s}}{\partial t} + (\mathbf{v}_{0s} + \mathbf{v}_{1s}) \cdot \nabla n_0 + \mathbf{v}_{0s} \cdot \nabla n_{1s} + (n_0 + n_{1s}) \nabla \cdot \mathbf{v}_{0s} + n_0 \nabla \cdot \mathbf{v}_{1s} = 0. \qquad (6.8)$$

This expression may be simplified immediately since several of its terms are identically zero. The plasma was prescribed to have a constant equilibrium number density which is equal for both species present so that $\nabla n_0 = 0$. Because the streaming velocities are aligned along the magnetic field, which lies in the $x-y$ plane perpendicular to its gradient, the terms involving $\nabla \cdot \mathbf{v}_{0s}$ also vanish and equation (6.8) reduces to

$$\frac{\partial n_{1s}}{\partial t} + \mathbf{v}_{0s} \cdot \nabla n_{1s} + n_0 \nabla \cdot \mathbf{v}_{1s} = 0.$$

None of the three remaining terms can be set automatically to zero since we have no information about the direction of oscillation of the first-order variables. In other words, this form of equation (6.8) is irreducible without further restrictions being placed on the model. It is, however desirable for us to achieve the maximum simplification from the outset. We wish to show that waves in non-uniform plasmas are not necessarily identical to those found under constant conditions. What better way to achieve this end than to demonstrate the difference between these two cases in a mathematically uncomplicated and "familiar" regime? By studying a specific case initially, it should also be possible to reduce the likelihood of error and to construct a test case for comparison with more general results.

To this end, we make a fundamental decision about the nature of the waves which we intend to study by assuming that there is no component of propagation in the $x-y$ plane ($k_x = k_y = 0$). Thus we may write $\nabla \equiv (0, 0, \partial/\partial z)$ acting on the perturbed quantities. This restriction allows us to eliminate a final term from equation (6.8), since $\mathbf{v}_{0s}$ is still perpendicular to ∇, even for first-order variables, and we have the final linearised form of equation (6.8):

$$\frac{\partial n_{1s}}{\partial t} + n_{0s} \nabla \cdot \mathbf{v}_{1s} = 0.$$

We are still interested in harmonically varying quantities at real frequencies and so perform the standard Fourier transform in time on all our linearised equations. Rearranging the above equation, so that the varying number density is expressed as a function of the perturbed velocity and known quantities, we obtain:

$$n_{1s} = \frac{n_0 v'_{sz}}{i\omega}.$$

4.2. Momentum equation

Care must be taken when linearising equations (6.1) to (6.5) because the extra non-zero equilibrium quantities occur in several places, giving rise to additional first-order terms which were not present in the equivalent equations of Chapter 5. Physically, the new equilibrium velocities couple new effects into the system which were previously second order and so were negligible under the assumption of linear wave propagation.

On linearising, equation (6.2) becomes:

$$m_s\left(\frac{\partial \mathbf{v}_{1s}}{\partial t} + (\mathbf{v}_{1s}\cdot\nabla)\mathbf{v}_{0s}\right) = q_s(\mathbf{E}_1 + \mathbf{v}_{0s}\times\mathbf{B}_1 + \mathbf{v}_{1s}\times\mathbf{B}_0)\ , \tag{6.9}$$

where the advective term, $\mathbf{v}_{0s}\cdot\nabla\mathbf{v}_{1s}$, has again been omitted from the total derivative because $\mathbf{v}_{0s}$ lies in the plane perpendicular to the direction of propagation.

Comparing equation (6.9) with its equivalent in the homogeneous model, equation (5.1), we note that the equilibrium flow has necessitated the inclusion of two additional terms. The first of these contains the derivative of the equilibrium velocity, which is a known function of the ambient magnetic field. The second, which involves the perturbed magnetic field, may be eliminated in favour of the perturbed electric field, by using the linearised form of Faraday's Law as shown below.

4.3. Maxwell's equations

The linearised Maxwell relations are unchanged from the homogeneous model and the linearised version of Faraday's law, equation (6.3) is:

$$\nabla\times\mathbf{E}_1 = -\frac{\partial \mathbf{B}_1}{\partial t}\ . \tag{6.10}$$

On assuming variation as $e^{-i\omega t}$ and expressing the perturbed magnetic field in terms of its electric counterpart, this may be rearranged to express $\mathbf{B}_1(\mathbf{E}_1)$:

$$\mathbf{B}_1 = \frac{1}{i\omega}(-E_{1y}', E_{1x}', 0). \tag{6.11}$$

The remaining Maxwell equation, which will be required in a subsequent section to derive the wave equation, is also unchanged in linearised form:

$$\nabla\times\mathbf{B}_1 = \mu_0 \, (\mathbf{J}_1 + \varepsilon_0 \frac{\partial \mathbf{E}_1}{\partial t}) \, ,$$

since $\nabla\times\mathbf{B}_0 = \mu_0\mathbf{J}_0$. After Fourier transforming to remove the time dependence, we have:

$$\nabla\times\mathbf{B}_1 = \mu_0 \, (\mathbf{J}_1 - \varepsilon_0 \, i \, \omega\mathbf{E}_1).$$

Note that we retain this form of the equation without simplification so that we can still write our wave equation (derived by taking the curl of equation (6.10)) in a style similar to the homogeneous wave equation, (5.5).

4.4. Current equation

The perturbed form of the current equation (6.5) has, like the momentum equation, gained a term which depends on the equilibrium velocity to become:

$$\mathbf{J}_1 = \sum_s q_s \, (n_0 \mathbf{v}_{1s} + n_{1s} \mathbf{v}_{0s} \,). \tag{6.12}$$

The first-order number density term in equation (6.12) is the reason for our earlier inclusion of the continuity equation to describe the density variation. We may use the result derived from equation (6.8), relating the perturbed number density and velocity, to eliminate the former from equation (6.12).

To simplify the notation considerably, we henceforth use subscripts only to distinguish between species and to denote equilibrium quantities. Thus any variable not subscripted zero will be assumed to be first order. We will also continue to use the conventions, $' \stackrel{\partial}{=} \partial/\partial z$ and $\dot{} \stackrel{\partial}{=} \partial/\partial t$.

5. The wave equation

We now have the basic elements required to form a complete wave equation in the electric field components. Having eliminated $\mathbf{B}$ and n_s, it only remains to express $\mathbf{v} = \mathbf{v}(\mathbf{E},\mathbf{E}')$, using equation (6.9), and hence $\mathbf{J} = \mathbf{J}(\mathbf{E},\mathbf{E}',\mathbf{E}'')$, using equation (6.12). This is the same procedure as was employed in Chapter 5 and we will again utilise the cyclotron and plasma frequencies in our matrix notation for simplicity. Substituting for the first-order magnetic field from equation (6.11) into the momentum equation, (6.9) written in tensor form:

$$\mathbf{M}\mathbf{v}_s = \frac{q_s}{m_s}\mathbf{E} + \frac{q_s}{m_s}\mathbf{N}\mathbf{E}',$$

which may be rearranged to give an expression for $\mathbf{v}_s$ in terms of $\mathbf{E}$:

$$\mathbf{v}_s = \frac{q_s}{m_s}\mathbf{M}^{-1}\mathbf{E} + \frac{q_s}{m_s}\mathbf{M}^{-1}\cdot\mathbf{N}\,\mathbf{E}'. \tag{6.13}$$

The tensors $\mathbf{M}$ and $\mathbf{N}$ introduced above are therefore given by:

$$\mathbf{M} = \begin{bmatrix} -i\omega & 0 & \varepsilon_s\Omega_{sy} + v_{0sx}' \\ 0 & -i\omega & -\varepsilon_s\Omega_{sy} + v_{0sy}' \\ -\varepsilon_s\Omega_{sy} & \varepsilon_s\Omega_{sx} & -i\omega \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \dfrac{v_{0sx}}{i\omega} & \dfrac{v_{0sy}}{i\omega} & 0 \end{bmatrix},$$

where we have used the shorthand $\Omega_{sx,y} = \varepsilon_s eB_{x,y}/m_s$. In the tensors, $\mathbf{M}$ and $\mathbf{N}$, the position dependence is contained in both the cyclotron frequencies and equilibrium velocities due to their variation with $\mathbf{B}_0$. The magnitudes of the tensor elements vary through the derivatives of ϕ, the angle which the background magnetic field makes with the x-axis, whilst their directions vary depending on the value of ϕ itself.

We see from equation (6.13) that we will need to determine the contents of the arrays $\mathbf{M}^{-1}$ and $\mathbf{M}^{-1}\cdot\mathbf{N}$ in order to express the velocity in terms of the electric field. The determinant of $\mathbf{M}$ is given by $i\omega(\omega^2-\Omega_s^2+\varepsilon_s\Omega_x v_{0sy}'-\varepsilon_s\Omega_y v_{0sx}')$. Writing $A_s = \begin{cases} \mu_0 n_0 e(1+m_e/m_i) & \text{for electrons} \\ -m_i/m_e\,\mu_0 n_0 e(1+m_e/m_i) & \text{ions} \end{cases}$ and using the results of §3.3, we find that

$$v_{0sx}' = \frac{(B_{0x}\phi')'}{A_s} = \frac{(-B_{0y}\phi'^2 + B_{0x}\phi'')}{A_s}$$

and

$$v_{0sy}' = \frac{(B_{0x}\phi'^2 + B_{0y}\phi'')}{A_s},$$

so that

$$\begin{aligned}\varepsilon_s(\Omega_x v_{0sy}' - \Omega_y v_{0sx}') &= \frac{\varepsilon_s e}{m_s A_s}(B_{0x}^2\phi'^2 + B_{0x}B_{0y}\phi'' + B_{0y}^2\phi'^2 - B_{0x}B_{0y}\phi'') \\ &= \frac{\varepsilon_s eB_0^2\phi'^2}{m_s A_s} \\ &= v_{0s}\varepsilon_s\Omega_s\phi',\end{aligned}$$

and thus one of the effects of the advective derivative is to shift the frequency given by the traditional denominator of the tensor elements all of which are derived from the determinant of **M**.

For convenience, we include at this stage two further matrices, derived from those above, which will be required shortly:

$$\mathbf{M}^{-1} = \frac{1}{\det}\begin{bmatrix} -\omega^2+\Omega_{sx}{}^2-\varepsilon_s\Omega_{sx}\mathrm{v}_{0sy}{}' & \varepsilon_s\Omega_{sx}(\varepsilon_s\Omega_{sy}+\mathrm{v}_{0sx}{}') & i\,\omega(\varepsilon_s\Omega_{sy}+\mathrm{v}_{0sx}{}') \\ \varepsilon_s\Omega_{sy}(\varepsilon_s\Omega_{sx}-\mathrm{v}_{0sy}{}') & -\omega^2+\Omega_{sy}{}^2+\varepsilon_s\Omega_y\mathrm{v}_{0sx}{}' & -i\,\omega(\varepsilon_s\Omega_{sx}-\mathrm{v}_{0sy}{}') \\ -i\,\omega\varepsilon_s\Omega_{sy} & i\,\omega\varepsilon_s\Omega_{sx} & -\omega^2 \end{bmatrix},$$

and

$$\mathbf{M}^{-1}\cdot\mathbf{N} = \frac{1}{\det}\begin{bmatrix} \mathrm{v}_{0sx}(\varepsilon_s\Omega_{sy}+\mathrm{v}_{0sx}{}') & \mathrm{v}_{0sy}(\varepsilon_s\Omega_{sy}+\mathrm{v}_{0sx}{}') & 0 \\ -\mathrm{v}_{0sx}(\varepsilon_s\Omega_{sx}-\mathrm{v}_{0sy}{}') & -\mathrm{v}_{0sy}(\varepsilon_s\Omega_{sx}-\mathrm{v}_{0sy}{}') & 0 \\ i\,\omega\mathrm{v}_{0sx} & i\,\omega\mathrm{v}_{0sy} & 0 \end{bmatrix},$$

where $\det = i\,\omega(\omega^2-\Omega_s{}^2+\mathrm{v}_{0s}\varepsilon_s\Omega_s\phi')$.

The final step which we must take before we can write down the wave equation is to express the current in terms of the electric field. We have already established from equation (6.8) that $n_s = \frac{n_0}{i\,\omega}\mathrm{v}_{sz}{}'$ and from equation (6.13) we know that the z component of the perturbed velocity is:

$$\mathrm{v}_{sz}{}' = \frac{q_s}{m_s}\Big[(\mathbf{M}^{-1})'\mathbf{E}+(\mathbf{M}^{-1}+(\mathbf{M}^{-1}\cdot\mathbf{N})')\mathbf{E}'+\mathbf{M}^{-1}\cdot\mathbf{N}\mathbf{E}''\Big]_z ,$$

where the subscript z indicates that only the third row of the tensor is being considered. Combining these two results and substituting in the linearised current equation, (6.12), enables us to write the perturbed current as a function of the electric field and its derivatives:

$$\mathbf{J} = \sum_s q_s(n_0\mathrm{v}_s + n_s\mathrm{v}_{0s})$$

$$= \sum_s q_s\,(n_0\mathrm{v}_s + \frac{n_0}{i\,\omega}\mathrm{v}_{sz}{}'\mathrm{v}_{0s})$$

$$= \sum_s q_s n_0(\frac{q_s}{m_s}\mathbf{M}^{-1}\mathbf{E} + \frac{q_s}{m_s}\mathbf{M}^{-1}\cdot\mathbf{N}\mathbf{E}')$$

$$+\sum_s \frac{q_s n_0}{i\,\omega}\mathrm{v}_{0s}\frac{q_s}{m_s}\Big[(\mathbf{M}^{-1})'\mathbf{E}+(\mathbf{M}^{-1}+(\mathbf{M}^{-1}\cdot\mathbf{N})')\mathbf{E}'+\mathbf{M}^{-1}\cdot\mathbf{N}\mathbf{E}''\Big]_z .$$

Gathering like terms, this may be written more concisely as:

$$\mathbf{J} = \sigma\cdot\mathbf{E} + \tau\cdot\mathbf{E}' + \rho\cdot\mathbf{E}''. \tag{6.14}$$

This should be contrasted with the equivalent expression for the case of a magnetic field making a fixed angle with the x-axis throughout the plasma, which was found to be $\mathbf{J} = \sigma\cdot\mathbf{E}$, where σ was identified with the conductivity tensor. In the present case however, it is impossible for us to relate the current density solely to the electric field. A similar result was found by Choudhury (1988), who formulated the electric displacement as

$$\mathbf{D} = \kappa\cdot\mathbf{E},$$

where the dielectric tensor κ consisted of a simple multiplicative part plus a differential operator component acting on the electric field.

We will now make a further decision regarding the variation with position of the zero-order magnetic field. We specify the functional form of ϕ to be linear with increasing z so that ϕ' becomes a constant and all higher derivatives vanish. In so doing, we introduce the rotational scalelength, l, defined by:

$$\phi = \frac{\pi z}{l},$$

thus ensuring that when $l \to \infty$, the gradient terms become negligible. The choice of such a simple magnetic field variation may seem at first a little restrictive but will enable us to produce an analytic solution. Even such an apparently minor deviation from the anisotropic, homogeneous cold plasma model will produce solutions of extremely different forms from those which would be predicted by the homogeneous dispersion relation. This should serve as a warning against relying on extensions of homogeneous theory to describe inhomogeneous plasmas. It will prove with hindsight that this particular variation is the only one which results in a readily soluble pair of coupled differential equations describing the electric field components. With our preliminary investigation completed, consideration may then be extended to more complex field structures where ϕ has other than linear dependence on z. The discussion of this greatly complicated case will be pursued more fully in Chapter 7.

We use the definitions of the tensors $\mathbf{M}$ and $\mathbf{N}$, $\mathbf{M}^{-1}$ and $\mathbf{M}^{-1}\mathbf{N}$, and extract a common factor of $n_{0s} q_s{}^2/m_s = \varepsilon_0 \omega_{ps}{}^2$ from equation (6.14) to write σ, τ and ρ as:

$$\sigma = \sum_s \frac{\varepsilon_0 \omega_{ps}^2}{i\,\omega(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s \Omega_s \phi')} \left[\mathbf{M}^{-1} + \frac{\phi'}{i\,\omega} v_{0s} : \mathbf{f}_1 \right],$$

$$\tau = \sum_s \frac{\varepsilon_0 \omega_{ps}{}^2}{i\,\omega(\omega^2 - \Omega_s^2 + v_{0s}\,\varepsilon_s\,\Omega_s\,\phi')} \left[\mathbf{M}^{-1}\mathbf{N} + \mathbf{v}_{0s} : \mathbf{f}_2 \right],$$

$$\rho = \sum_s \frac{\varepsilon_0 \omega_{ps}{}^2}{i\,\omega(\omega^2 - \Omega_s^2 + v_{0s}\,\varepsilon_s\,\Omega_s\,\phi')} \mathbf{v}_{0s} : \mathbf{v}_{0s},$$

where $\mathbf{f}_1 = (-i\,\omega\varepsilon_s\,\Omega_{sx}, -i\,\omega\varepsilon_s\,\Omega_{sy}, 0)$ and $\mathbf{f}_2 = (-\varepsilon_s\,\Omega_{sy} - \phi' v_{0sy}, \varepsilon_s\,\Omega_{sx} + \phi' v_{0sx}, i\,\omega)$ and : indicates an outer product.

Having now reduced our description of the model to this form, with all the variables related to the oscillating electric field and its derivatives, we may proceed to generate the wave equation in an analogous manner to Chapter 5:

$$\begin{aligned} \nabla\times(\nabla\times\mathbf{E}) &= -\nabla\times\dot{\mathbf{B}} \\ &= i\,\omega\,\mu_0(\mathbf{J} - i\,\omega\,\varepsilon_0\mathbf{E}) \\ &= \frac{\omega^2}{c^2}\,(\mathbf{I} + \frac{i\,\sigma}{\varepsilon_0\omega})\cdot\mathbf{E} + i\,\omega\,\mu_0(\tau\cdot\mathbf{E}' + \rho\cdot\mathbf{E}'') \\ &= \frac{\omega^2}{c^2}\,\varepsilon\cdot\mathbf{E} + i\,\omega\,\mu_0(\tau\cdot\mathbf{E}' + \rho\cdot\mathbf{E}''), \end{aligned} \tag{6.15}$$

where ε plays the role of the conventional dielectric tensor and τ and ρ which multiply the spatial derivatives of the electric field have been defined above. It should be noted that the substantially different nature of the Ohm's Law, equation (6.14), which applies here has led to an equivalent alteration of the wave equation. In addition to the two completely new tensors τ and ρ, even the elements of the dielectric tensor, ε, are *not* those of Chapter 5 but are shown below to be considerably modified by the gradients in the equilibrium quantities.

6. Tensor definitions

6.1. The modified dielectric tensor

By inspection of equation (6.15) and from the the conductivity tensor given in the previous section, the dielectric tensor may be written, using notation analogous to that of ordinary modes, sums and differences from the homogeneous case:

$$\varepsilon = \begin{bmatrix} (P - \phi' T_1/\omega)\mu^2 + \tilde{S}\nu^2 & (P - \phi' T_1/\omega - \tilde{S})\mu\nu & i\,\nu(\tilde{D} - \phi' T_2/\omega) \\ (P - \phi' T_1/\omega - \tilde{S})\mu\nu & (P - \phi' T_1/\omega)\nu^2 + \tilde{S}\mu^2 & -i\,\mu(\tilde{D} - \phi' T_2/\omega) \\ -i\tilde{D}\,\nu & i\tilde{D}\,\mu & \tilde{S} \end{bmatrix},$$

where $P = 1 - \sum_s \omega_{ps}^2/\omega^2$ retains its definition from Chapter 5 but, because of the alteration in the denominator of the tensor elements, the definitions of S and D, are modified to:

$$\tilde{S} = 1 - \sum_s \frac{\omega_{ps}^2}{(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s\Omega_s\phi')} \quad \text{and} \quad \tilde{D} = \sum_s \frac{\omega_{ps}^2\varepsilon_s\Omega_s}{\omega(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s\Omega_s\phi')}.$$

The terms containing the equilibrium velocities in their numerators have been combined into the new quantity T_1:

$$T_1 = \sum_s \frac{\omega_{ps}^2\varepsilon_s\Omega_s v_{0s}}{\omega(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s\Omega_s\phi')}.$$

6.2. The tensors, τ and ρ

Two similar quantities may be used in the expressions for the remaining tensors. We define:

$$T_2 = \sum_s \frac{\omega_{ps}{}^2 v_{0s}}{(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s\Omega_s\phi')} \quad \text{and} \quad T_3 = \sum_s \frac{\omega_{ps}{}^2 v_{0s}^2}{\omega(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s\Omega_s\phi')}.$$

Although the T terms appear to share several common factors they are not in fact multiples of each other because they are all sums over the species present.

τ and ρ thus simplify in terms of these new tensor elements to

$$\tau = \frac{\varepsilon_0}{i}\begin{bmatrix} -2\phi' T_3\mu\nu & T_1 + \phi' T_3(\mu^2 - \nu^2) & i\mu T_2 \\ -T_1 + \phi' T_3(\mu^2 - \nu^2) & 2\phi' T_3\mu\nu & i\nu T_2 \\ i\mu T_2 & i\nu T_2 & 0 \end{bmatrix},$$

and

$$\rho = \frac{\varepsilon_0}{i} T_3 \begin{bmatrix} \mu^2 & \mu\nu & 0 \\ \mu\nu & \nu^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We have now extended our examination of waves in a "cold plasma" with a spatially rotating field far beyond the limiting case of the uniform field discussed in Chapter 5. This special case may, however, be used to test our analysis thus far. In the limit of $\phi \rightarrow$ constant, (i.e. removing the position dependence) we should be able to reproduce exactly the results of the homogeneous case presented in Chapter 5.

$\phi =$ constant implies $\phi' = 0$ and so v_{0s} must vanish by definition. As a direct consequence, $T_1 = T_2 = T_3 = 0$ and $\tilde{S}$ and $\tilde{D}$ reduce to their homogeneous counterparts, S and D. Thus the tensors τ and ρ vanish as do those elements of the dielectric tensor which involve gradient terms through T_1, leaving the Ohm's law in its original form, $\mathbf{J} = \sigma\cdot\mathbf{E}$.

To complete the comparison with the inhomogeneous case contained in the previous chapter, the magnetic field must be aligned along one axis. Here, we have chosen the variation to be in the z-direction and so the magnetic field lies in the $x-y$ plane. It is therefore convenient to select the x-axis as our single preferred direction in the homogeneous limit whereas the magnetic field lines of Chapter 5 were parallel to the z-axis. We thus require to perform a simple rotational transformation on the dielectric tensor of Chapter 5 to demonstrate that it is identical to that generated by substituting $\mathbf{B}_0 = B_0\hat{\mathbf{x}}$ (so that $\mu = 1$ and $\nu = 0$) in ε above:

$$\varepsilon = \begin{bmatrix} P & 0 & 0 \\ 0 & S & -iD \\ 0 & iD & S \end{bmatrix}.$$

The wave equation, (6.15), reduces to

$$\nabla\times\nabla\times\mathbf{E} - \frac{\omega^2}{c^2}\varepsilon\cdot\mathbf{E} = 0, \tag{5.5}$$

in full agreement with the expected behaviour (Stix, 1962). The general case of non-constant magnetic field angle therefore reduces to the predicted result for a special case and so we may continue with its investigation with confidence.

7. Coupled equations

The standard way to solve a system such as equation (6.15) would be to separate it into its individual component equations and then eliminate two of the dependent variables. This would leave a single ordinary differential equation in E_x, say, which could then be solved, with its solutions giving the natural modes of the system. This is unfortunately not as easy as it sounds in this case. On eliminating E_z, which is a straightforward procedure performed below, we are left with two coupled second order differential equations in E_x and E_y. Both of these equations have periodic and variable coefficients due to their dependence on the equilibrium quantities and, although combining them into a single fourth order differential equation is feasible, it is of no advantage. The fourth order equation resulting from such an elimination procedure would have extremely complicated coefficients since the trigonometric factors would not be lost after successive differentiation in the way that polynomial coefficients might. Hence we would be left with a fourth order differential equation with periodic coefficients for which there is no standard method of solution. Since one of our fundamental aims is to produce an analytic solution which we may compare with that for the non-rotating model, we do not wish to be forced to resort to numerical methods of solution. We therefore seek an alternative approach to solving our pair of coupled differential equations in E_x and E_y.

As stated above, E_z may be eliminated quite easily from the system. The third component of the vector equation (6.15) contains no derivatives of E_z so that it may be eliminated in favour of E_x, E_y and their first derivatives:

$$E_z = -\frac{1}{\varepsilon_{33}}\left(\frac{1}{\omega}(\tau_{31}E_x{}' + \tau_{32}E_y{}') + \varepsilon_{31}E_x + \varepsilon_{32}E_y\right),$$

where we have removed the factor ε_0/i from the elements of the matrices τ and ρ. Substituting this expression for E_z along with its derivative into the remaining two components of equation (6.15) yields a pair of coupled equations in E_x and E_y. At this point, we introduce three new quantities to ease notation. These quantities are designed to collect together the occurrences of the T tensor elements, containing the bulk of the information about variation with ϕ'. Their precise definitions and a discussion of their functional dependence is contained in the following section.

We may now use the expression for E_z to eliminate it from our system, leaving us with a pair of coupled differential equations for E_x and E_y:

$$(1+\psi\mu^2)E_x{}'' - 2\phi'\psi\mu\nu E_x{}' + \left(\frac{\omega^2}{c^2}P\mu^2 + \frac{\omega^2}{c^2}\frac{\tilde{R}\tilde{L}}{\tilde{S}}\nu^2 - \phi'\chi\mu^2 + \Xi\nu^2\right)E_x$$

$$= -\psi\mu\nu E_y{}'' - (\chi + \phi'\psi(\mu^2-\nu^2))E_y{}' - \mu\nu\left(\frac{\omega^2}{c^2}(P - \frac{\tilde{R}\tilde{L}}{\tilde{S}} - \phi'\chi - \Xi\right)E_y \quad (6.16)$$

and

$$(1+\psi\nu^2)E_y{}'' + 2\phi'\psi\mu\nu E_y{}' + \left(\frac{\omega^2}{c^2}P\nu^2 + \frac{\omega^2}{c^2}\frac{\tilde{R}\tilde{L}}{\tilde{S}}\mu^2 - \phi'\chi\nu^2 + \Xi\mu^2\right)E_y$$

$$= -\psi\mu\nu E_x{}'' + (\chi - \phi'\psi(\mu^2-\nu^2))E_x{}' - \mu\nu\left(\frac{\omega^2}{c^2}(P - \frac{\tilde{R}\tilde{L}}{\tilde{S}}) - \phi'\chi - \Xi\right)E_x, \quad (6.17)$$

where we have introduced the quantities $\tilde{R} = \tilde{S} + \tilde{D}$ and $\tilde{L} = \tilde{S} - \tilde{D}$ so that the notation matches that of Chapter 5 as closely as possible. These quantities do not, however, represent the natural modes of the system for oscillations parallel to the magnetic field in the way that their counterparts in the homogeneous case described the right and left circularly polarised waves. Because P is independent of the magnetic field, its expression has not been altered.

It is at this stage that the basic difference between Choudhury's philosophy and ours forces the analyses irreconcilably apart. Choudhury's aim is to examine the response of a specific device (a tokamak) to a specific input frequency (the lower hybrid) and he may thus eliminate terms as it becomes apparent from additional experimental information that they are unimportant. He thus sets $E_z = 0$, arguing that the high conductivity in the plasma shorts out the parallel electric field. This approximation leads to a considerable simplification in the resulting equations so that the final system is second order - equivalent to only the 'fast' mode being considered. Although his original magnetic field consists of both toroidal and poloidal components, only leading order terms in the poloidal field are retained because it is known to contribute only a small fraction to the total magnetic field strength. Similarly, the fact that the parallel wavenumber is approximately constant allows him to solve an eigenvalue problem for the perpendicular wavenumber only. This single-minded attack on a known objective is particularly appropriate in application to a particular experiment but our approach is somewhat different. We wish to examine the consequences of introducing position-dependence into a well-understood situation. We are not constrained by the specifications of any apparatus but must retain a self-consistency in the equations describing the model. We are interested in generating an analytic solution to the complete problem and consider any effect arising, however insignificant it may at first appear. From this general picture, it should be possible to take an overall view of the consequences that inhomogeneity may have for wave propagation and make more general predictions about the changes in behaviour of the modes.

7.1. χ, ψ and Ξ

7.1.1. χ and ψ

The new quantities, χ and ψ which were introduced above are defined by:

$$\chi = \frac{\omega}{c^2}(\mathrm{T}_1 + \frac{\tilde{D}}{\tilde{S}}\mathrm{T}_2) \quad \text{and} \quad \psi = \frac{\omega}{c^2}(\mathrm{T}_3 + \frac{\mathrm{T}_2^2}{\tilde{S}\,\omega}).$$

Both ψ and χ are functions of the frequency and we apply a similar analysis to them as was used to investigate the variation with ω of the ordinary and extraordinary modes in Chapter 5. We begin by rewriting them in terms of the non-dimensional quantities, X and Y, which are themselves defined in terms of the *electron* cyclotron and plasma frequencies plus Z which incorporates the derivative terms:

$$X = \frac{\omega}{\omega_{pe}} \qquad Y = \frac{\Omega_e}{\omega_{pe}} \qquad Z = \frac{v_{0e}\,\phi'}{\omega_{pe}}.$$

On substituting these dimensionless quantities into the definitions of T_1, T_2 and T_3 derived earlier, we establish that:

$$T_1 = -\frac{\nu_{0e} Y(1+r)(X^2(1-r+r^2) - r^2 Y(Y+Z))}{X(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))},$$

$$T_2 = \frac{\nu_{0e} X^2(1-r^2)}{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))}$$

and

$$\omega T_3 = \frac{\nu_{0e}^2 (1+r)(X^2(1-r+r^2) - r^2 Y(Y+Z))}{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))},$$

where r is the electron to ion mass ratio. Similar expressions to those for S and D, derived in Chapter 5, may be found for $\tilde{S}$ and $\tilde{D}$:

$$\tilde{S} = \frac{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z)) - (1+r)(X^2 - rY(Y+Z))}{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))}$$

and

$$\tilde{D} = \frac{XY(r^2-1)}{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))}.$$

Before constructing χ and ψ from these, we will compare the relative sizes of the terms containing Z with the other terms in the denominators of the above expressions.

$$\left| \frac{Z}{Y} \right| = \frac{\nu_{0e}\phi'}{\Omega_e} \approx \frac{\phi'^2 m_e}{\mu_0 n_0 e^2} = \left(\frac{\phi' c}{\omega_{pe}}\right)^2.$$

For a typical thermonuclear plasma $n = 10^{19}\mathrm{m}^{-3}$ so that $\left|\frac{Z}{Y}\right| \approx 2\times 10^{-6}\phi'^2$. Thus, for a plasma with this density, Z will be less than 10% of Y as long as $\phi' \gtrsim 220$, i.e. the scalelength is greater than 0.014 m.

Substituting the dimensionless forms of T_1 etc. into the definitions of χ and ψ above yields:

$$\chi = \frac{\omega}{c^2} \frac{\nu_{0e} Y(1+r)(r^2 Y(Y+Z) - X^2(1-r+r^2) + r(1+r))}{X((X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z) - (1+r)(X^2 - rY(Y+Z)))}, \quad (6.18)$$

and

$$\psi = -\frac{1}{c^2} \frac{\nu_{0e}^2 (1+r)(r^2 Y(Y+Z) - X^2(1-r+r^2) + r(1+r))}{((X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z)) - (1+r)(X^2 - rY(Y+Z)))}.$$

In this form, it is obvious that ψ and χ are far from being independent of each other. ψ is in fact a multiple of χ, where the multiplicative factor is frequency-independent. This is a great boon to us, since calculating their ratio and defining it to be a new quantity enables us to distinguish between a number of independent factors influencing the behaviour of equations (6.16) and (6.17). The required ratio is given by:

$$\frac{\psi}{\chi} = -\frac{v_{0e}X}{\omega Y} = -\frac{v_{0e}}{\Omega_e} = \frac{v_{0i}}{\Omega_i},$$

where we have returned from the dimensionless notation to the physically more significant quantities of velocity and gyrofrequency. We define $\Lambda = v_{0e}/\Omega_e$. Note that there is a change of sign between the two equivalent expressions for Λ. In addition to the scale factor r which transforms the electron cyclotron frequency and equilibrium velocity to their counterparts for the ions, there is an extra sign factor due to the electrons and ions having opposite streaming directions in equilibrium.

From our earlier definitions of equilibrium velocity and cyclotron frequency in terms of the basic plasma parameters such as ambient magnetic field strength and neutral number density, we see that Λ is independent of $\mathbf{B}_0$ (as well as frequency). Finally, we express Λ in its fundamental form:

$$\Lambda = \frac{m_e m_i}{\mu_0 e^2(m_e + m_i)} \frac{\phi'}{n_0}.$$

Thus, once the background number density of the plasma is known, Λ is completely determined by how tightly the magnetic field "winds around" the vertical axis. Therefore when ϕ is directly proportional to distance ($\phi' =$ constant), Λ is simply a constant which multiplies several terms involving χ in our coupled equations. Similarly, $\Lambda = 0$ for a magnetised plasma only when the plasma is homogeneous.

The behaviour of ψ or χ is rather more complicated than that of Λ because of their variation with frequency and so is more interesting. Having demonstrated the linear relationship between ψ and χ, it is sufficient for us to examine the frequency-dependence of only one of these and in what follows we choose to consider χ.

First, we will tackle the asymptotic values of χ, then progress to its zeros and singularities. For very large values of X (i.e at high frequencies), we see from equation (6.18) that $\chi \rightarrow 0$. At the opposite end of the frequency spectrum, as $X \rightarrow 0$, we find that, dropping factors of r with respect to 1:

$$\chi \rightarrow \frac{v_{0e}\,\omega_{pe}}{c^2(Y+Z)}$$

$$\approx \frac{\omega_{pe} v_{0e}}{c^2 Y}$$

$$= \frac{B_0 \phi' m_e n_0 e^2}{\mu_0 n_0 e c^2 \varepsilon_0 m_e e B_0}$$

$$= \phi',$$

which is a refreshingly simple result after the complicated nature of the original expression. It seems physically plausible, particularly at low frequencies, that the quantities introduced by the inhomogeneity in the magnetic field should be related so simply to the field's degree of rotation.

Again using the approximation $r \ll 1$, the denominator of χ may be written:

$$X^4 - X^2(1+Y(Y+Z)) + rY(Y+Z)(1+rY(Y+Z)),$$

so that the resonances of χ are given by

$$2X^2 = 1+Y(Y+Z) \pm (1+Y(Y+Z))(1 - \frac{4rY(Y+Z)(1+rY(Y+Z))}{(1+Y(Y+Z))^2})^{1/2},$$

which may be approximated using the binomial expansion as:

$$X^2 = 1+Y(Y+Z) \quad , \quad X^2 = \frac{rY(Y+Z)(1+rY(Y+Z))}{(1+Y(Y+Z))}.$$

The frequencies defined by these expressions are very close to the upper and lower hybrid frequencies respectively and in the limit $Z \rightarrow 0$, χ will share the resonances of the extraordinary mode which were shown in Chapter 5 to be:

$$\omega_{uh}^2 = \Omega_e^2 + \omega_{pe}^2 \quad \text{and} \quad \omega_{lh}^2 = \frac{\Omega_e \Omega_i (\omega_{pe}^2 + \Omega_e \Omega_i)}{\omega_{pe}^2 + \Omega_e^2}.$$

Figure (6.2) shows schematically the variation of χ with ω while Figure (6.3) illustrates that the upper hybrid resonance is indeed a good approximation to one resonance of χ.

The zero of χ occurs when

$$X^2 \approx r^2 Y^2 + r,$$

which may be rewritten in terms of the intrinsic plasma frequencies as:

$$\omega^2 = \omega_{pi}^2 + \Omega_i^2.$$

Notice that this is of an analogous form to the upper hybrid frequency which pertained to electron motion. As a numerical example, consider $B_0 = 0.1\text{T}$ and

$n_0 = 10^{19}m^{-3}$. The ion cyclotron frequency is $9.58\times10^6 s^{-1}$ whereas the ion plasma frequency, of $4.16\times10^9 s^{-1}$, is 3 orders of magnitude greater and so $\omega_{pi} = 0$ is almost identical to $\chi = 0$. For these parameters, $\omega_{lh} \approx 4.08\times10^8 s^{-1}$ and $\omega_{uh} \approx 1.79\times10^{11} s^{-1}$.

7.1.2. Ξ

The final new variable introduced in equations (6.16) and (6.17) is defined by:

$$\Xi = \frac{\omega}{c^2}\frac{\tilde{D}}{\tilde{S}}\mathrm{T}_2\phi'$$

$$= -\frac{\nu_{0e}X^3Y(r^2-1)^2}{\mathrm{num}\tilde{S}\,\det\tilde{S}},$$

where $\mathrm{num}\tilde{S}$ and $\det\tilde{S}$ denote the numerator and denominator of $\tilde{S}$ respectively. It is clear from this expression that Ξ has no physically significant zeros but has, however, four resonances. From our previous analysis, we recognise that these resonances will occur near Ω_i, Ω_e, ω_{lh} and ω_{uh}, i.e. both cyclotron frequencies as well as the lower and upper hybrid frequencies respectively. As $X \to 0$ and as $X \to \infty$, $\Xi \to 0$ from below. The resonances of $\Xi(\omega)$ are illustrated in Figure (6.4).

It is also of interest to consider under what circumstances Ξ and χ are equal, if at all. Non-dimensionalising using ϕ', gives $\chi_0 = \chi/\phi'$ and $\Xi_0 = \Xi/\phi'^2$. Now

$$\chi_0 - \Xi_0 = \frac{\omega}{\phi' c^2}(\mathrm{T}_1 + \frac{\tilde{D}}{\tilde{S}}\mathrm{T}_2) - \frac{\omega\phi'}{c^2}\frac{\tilde{D}\mathrm{T}_2}{\tilde{S}\phi'^2} = \frac{\omega \mathrm{T}_1}{\phi' c^2}.$$

Using the definition of T_1 from section §7.1.1, we see that the function $\chi_0 - \Xi_0$ has resonances near the ion and electron cyclotron frequencies and a zero where $X^2(1-r+r^2) = r^2Y(Y+Z)$. For $r \ll 1$ and $Z \ll Y$, the latter expression simplifies to $X^2 = r^2Y^2$ which represents the ion cyclotron frequency. $\chi_0 - \Xi_0$ therefore has both a zero and a resonance situated close together in the vicinity of the ion cyclotron frequency. (The resonance is due to Ξ_0 having a resonance in this region while χ_0 does not.) Although χ_0 and Ξ_0 are closely related and indeed coincide at a specific frequency, it is equally possible for χ_0 to vanish when Ξ_0 does not, or for Ξ_0 to tend to infinity when χ_0 does not. We will use this information later when χ_0 and Ξ_0 are required as input to determine the number of propagating modes in the system.

7.1.3. $\tilde{R}\tilde{L}/\tilde{S}$

One variable which remains to be discussed is the one which is analogous to the extraordinary mode. It was remarked in Chapter 5 that although the right and left circularly polarised waves (R and L) had resonances at the ion and electron cyclotron frequencies, the extraordinary mode (RL/S) did not because of a cancellation with identical terms in S. The situation has now been altered because of the introduction, due to the advective derivative, of the term $\varepsilon_s \Omega_s v_{0s} \phi'$ in the denominators of the dielectric tensor components. As a result, the cancellation which occurred before no longer holds and two additional zeros, each near a new resonance, appear in $\tilde{R}\tilde{L}/\tilde{S}(\omega)$. None of these four features were present in $RL/S(\omega)$. It should be borne in mind that the quantities $\tilde{R}$, $\tilde{L}$ and $\tilde{S}$ do *not* represent the natural modes in the inhomogeneous plasma but are simply convenient extensions of our previous notation.

$\tilde{S}$ was defined in a previous section but will again prove most useful in our X, Y, Z notation:

$$\tilde{S} = 1 - \sum_s \frac{\omega_{ps}^2}{(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s \Omega_s \phi')}$$

$$= \frac{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z)) - (1+r)(X^2 - rY(Y+Z))}{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))}.$$

Expressions for $\tilde{R}$ and $\tilde{L}$ may be calculated from $\tilde{S}$ and $\tilde{D}$ to be:

$$\tilde{R}, \tilde{L} = 1 - \sum_s \frac{\omega_{ps}^2 (\omega \mp \varepsilon_s \Omega_s)}{\omega(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s \Omega_s \phi')}$$

$$= 1 - (1+r)\frac{X^2 \pm XY(1-r) - rY(Y+Z)}{(X^2 - Y(Y+Z))(X^2 - r^2 Y(Y+Z))}.$$

Taking all terms over the common denominator, we see that:

$$\frac{\tilde{R}\tilde{L}}{\tilde{S}} = \frac{\mathrm{num}(\tilde{R})\,\mathrm{num}(\tilde{L})}{\mathrm{num}(\tilde{S})\,\mathrm{den}(\tilde{S})},$$

which shows that the zeros of the function occur when the numerator of $\tilde{R}$ or $\tilde{L}$ goes to zero and that the resonances will be found when either the numerator or denominator of $\tilde{S}$ vanishes. We have seen in the preceding sections that $\mathrm{num}(\tilde{S}) \approx 0$ close to the upper and lower hybrid resonances. Also $\mathrm{den}(\tilde{S}) = 0$ is satisfied when

$$X^2 = Y(Y+Z) \quad \text{or} \quad X^2 = r^2 Y(Y+Z),$$

which represent

$$\omega^2 = \Omega_e^2 + \Omega_e v_{0e} \phi' \quad \text{and} \quad \omega^2 = \Omega_i^2 - \Omega_i v_{0i} \phi'.$$

Thus the remaining resonances of $\tilde{R}\tilde{L}/\tilde{S}$ occur very slightly above the ion and electron cyclotron frequencies. (Note that these frequencies are higher than the cyclotron frequencies since we have taken the direction in which the electrons flow in equilibrium to be positive and so both v_{0e} and $-v_{0i}$ are positive). The zeros of $\tilde{R}$ are given by

$$(X^2-Y(Y+Z))(X^2-r^2Y(Y+Z)) - (1+r)(X^2+XY(1-r)-rY(Y+Z)) = 0,$$

which only factors easily under the approximation $|Z| \ll |Y|$, giving

$$(X-Y)(X+Y)(X-rY)(X+rY) - (1+r)(X-rY)(X+Y) = 0.$$

Since $r \ll 1$, we may reduce this further to:

$$(X+Y)(X-rY)(X^2-XY-rY^2-1) = 0,$$

with real zeros at $X = rY$ and the solution of $X^2-XY-rY^2-1 = 0$, i.e. $\omega = \Omega_i$ and $\omega = \omega_R$ which is what we would have expected from Chapter 5. Similarly, $\tilde{L}$ has zeros near Ω_e and ω_L. The zeros at the ion and electron cyclotron frequencies are not coincident with the respective resonances but occur very close by.

The asymptotic behaviour of $\tilde{R}\tilde{L}/\tilde{S}$ is similar to that of RL/S. As $X \to \infty$, $\tilde{R}\tilde{L}/\tilde{S} \to 1$ and as $X \to 0$,

$$\frac{\tilde{R}\tilde{L}}{\tilde{S}} \to \frac{rY(Y+Z)+1+r}{rY(Y+Z)}.$$

Thus for high frequencies, the behaviour of the extraordinary mode and the new variation are identical, the most striking differences occurring in two small regions near the cyclotron frequencies. This can best be appreciated by comparing the schematic diagrams of Figures (6.5) and (5.3).

8. Solution of the coupled equations

This examination of the variation of χ and $\tilde{R}\tilde{L}/\tilde{S}$ with frequency, combined with our discussion of the dependences of Λ, completes the description of the behaviour of the coefficients in our pair of coupled equations. Λ clearly reflects the degree of rotation of the magnetic field whilst the remaining quantities contain in addition more complicated dependences on the ambient density, magnetic field strength and wave frequency and are not directly proportional to a single parameter.

Since there is no need to write it out explicitly, the factor ω^2/c^2 will henceforth be absorbed into the dielectric tensor components P, $\tilde{D}$, $\tilde{R}$, $\tilde{L}$ and $\tilde{S}$. We proceed by expressing all occurrences of ψ as the product of χ and $-\Lambda$ and rewrite equations (6.16) and (6.17) accordingly:

$$(1-\Lambda\chi\mu^2)E_x''+2\phi'\Lambda\chi\mu\nu E_x'+(P\mu^2+\frac{\tilde{R}\tilde{L}}{\tilde{S}}\nu^2-\phi'\chi\mu^2+\Xi\nu^2)E_x$$

$$=\Lambda\chi\mu\nu E_y''-(\chi-\phi'\Lambda\chi(\mu^2-\nu^2)E_y'-\mu\nu(P-\frac{\tilde{R}\tilde{L}}{\tilde{S}}-\phi'\chi-\Xi)E_y \quad (6.19)$$

and

$$(1-\Lambda\chi\nu^2)E_y''-2\mu\nu\phi'\Lambda\chi E_y'+(P\nu^2+\frac{\tilde{R}\tilde{L}}{\tilde{S}}\mu^2-\phi'\chi\nu^2+\Xi\mu^2)E_y$$

$$=\Lambda\chi\mu\nu E_x''+(\chi+\phi'\Lambda\chi(\mu^2-\nu^2))E_x'-\mu\nu(P-\frac{\tilde{R}\tilde{L}}{\tilde{S}}-\phi'\chi-\Xi)E_x\,.(6.20)$$

8.1. Bloch's theorem

Before calculating explicitly the solutions for E_x and E_y from equations (6.19) and (6.20), we will predict the general form of these solutions using information from other disciplines. Differential equations with periodic coefficients have been studied in other contexts for many years, although few have been completely solved. A close analogy to the current problem is that of waves propagating in a three dimensional lattice, which has long interested solid state physicists (Brillouin, 1953). The equation for wave propagation in a one-dimensional continuous medium may be written:

$$\frac{\partial^2\psi}{\partial x^2}+F(x)\psi=0.$$

When the function F, which has period π in x, contains only a single cosine term, this equation becomes Mathieu's equation (Mathieu, 1868; McLachlan, 1947). (Mathieu's equation may be obtained from the elliptic cylinder equation by transformation of the independent quantity $z\rightarrow\cos\phi$.) Floquet (1883) discovered that the general solution of Mathieu's equation could be written as $\Psi=e^{i\mu x}A(x)$ where $A(x)$ is also a π-periodic function and μ is a constant called the characteristic exponent (Whittaker and Watson, 1927). In practice, calculating the two values of μ which determine the two independent solutions is an extremely arduous task with special cases leading to the generation of a range of special functions also called after Mathieu. In one method of solution, a trial solution in the form of a Laurent series, $\sum_{n=-\infty}^{\infty} A_n e^{i(\mu+2n)x}$, is substituted into the differential equation to yield a recursion relation for the coefficients. The exponent μ and the coefficients in the Fourier expansion may be calculated using continued fractions or from the solution of an infinite determinant. (Often, solution of this infinite determinant is facilitated by using physical arguments to set the majority of its entries to zero.)

In application to solid state physics, Felix Bloch (1928) proved the important extension of Floquet's theorem to three dimensions, viz. that the solutions of the Schrödinger equation for a periodic potential must be of the form:

$$\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) e^{i \mathbf{k} \cdot \mathbf{r}},$$

where $u_{\mathbf{k}}(\mathbf{r})$ has the period of the crystal lattice.

Thus by analogy, we anticipate that the solution which we are seeking should be a travelling wave, modulated by a function with the periodicity of the medium through which it is propagating - arising from the periodic structure of the magnetic field.

8.2. Transformation of reference frame

In an earlier section, we argued the case for avoiding the formation of a fourth-order equation with periodic coefficients. Our policy must therefore be to find a way of solving equations (6.19) and (6.20) together. These coupled equations exhibit a marked symmetry and it is this property which we will exploit in producing solutions. As we move between equations (6.19) and (6.20), we see that there is an interchanging of the direction cosines, μ^2 becomes ν^2 (with sign changes in the coefficients of the first derivatives) and the products $\mu\nu$ still multiply the same coefficients except for one change of sign. Also the right hand sides of both equation, which can be considered to be coupling terms, are of very similar forms.

Bearing in mind this symmetry, we introduce the variables

$$E_+ = E_x + i\, E_y \quad \text{and} \quad E_- = E_x - i\, E_y, \tag{6.21}$$

which are complex conjugates and therefore contain a significant degree of symmetry themselves. A similar technique is used in quantum mechanics where the angular momentum operators $\hat{l}_+$ and $\hat{l}_-$ are analogously defined from the components of the total angular momentum (Cassels, 1982). E_+ and E_- are useful to us because they may be thought of as "following" the rotation of the field. In fact, this transformation of reference frame is probably the single most crucial factor in allowing us to find a simple analytic solution. In turn, this transformation is only useful because of the high degree of symmetry in the system and would not prove valuable in a much more general case.

The quantities P and $\tilde{R}\tilde{L}/\tilde{S}$ will not often occur singly in further analysis but as part of their sum and difference and so we introduce the notation $\kappa = \frac{1}{2}(P + \tilde{R}\tilde{L}/\tilde{S})$ and $\delta = \frac{1}{2}(P - \tilde{R}\tilde{L}/\tilde{S})$.

We now derive a new pair of equations from (6.19) and (6.20) expressed in terms of the new dependent variables E_+ and E_-. This set is also a pair of coupled, second order ordinary differential equations but has several advantages over the original pair. No information is lost in this process. The two sets of equations are equivalent and solutions for either two dependent variables instantly provide the solutions for the alternative ones via algebraic manipulation of the definitions in equation (6.21).

In what follows, we use the familiar trigonometric identities:

$$\cos^2\phi = \tfrac{1}{2}(1+\cos2\phi) \qquad \sin^2\phi = \tfrac{1}{2}(1-\cos2\phi) \qquad \sin\phi\cos\phi = \tfrac{1}{2}\sin2\phi,$$

and employ complex exponential notation, $e^{\pm 2i\phi} = \cos2\phi \pm i\sin2\phi$.

To construct the equations in E_+ and E_-, we add equation (6.19) to the product of i and equation (6.20) yielding:

$$E_+'' - \frac{\Lambda\chi}{2}(E_+'' + \cos2\phi E_-'') + \phi'\Lambda\chi\sin2\phi E_-' + \tfrac{1}{2}(P - \phi'\chi)(E_+ + \cos2\phi E_-)$$

$$+ \tfrac{1}{2}(\frac{\tilde{R}\tilde{L}}{\tilde{S}} + \Xi)(E_+ - \cos2\phi E_-)$$

$$= i\frac{\Lambda\chi}{2}\sin2\phi E_-'' + i\chi E_+' + i\phi'\Lambda\chi\cos2\phi E_-' - \tfrac{1}{2}i(P - \frac{\tilde{R}\tilde{L}}{\tilde{S}} - \phi'\chi - \Xi)\sin2\phi E_-,$$

which gives finally:

$$(1 - \frac{\Lambda\chi}{2})E_+'' - i\chi E_+' + (\kappa - \frac{\phi'\chi}{2} + \frac{\Xi}{2})E_+$$

$$= (\frac{\Lambda\chi}{2}E_-'' + i\frac{\phi'\Lambda\chi}{2}E_-' - (\delta - \frac{\phi'\chi}{2} - \frac{\Xi}{2})E_-)e^{2i\phi}. \qquad (6.22)$$

To generate our second equation in E_+ and E_-, we subtract i times equation (6.20) from equation (6.19) to obtain:

$$(1 - \frac{\Lambda\chi}{2})E_-'' + i\chi E_-' + (\kappa - \frac{\phi'\chi}{2} + \frac{\Xi}{2})E_-$$

$$= (\frac{\Lambda\chi}{2}E_+'' - i\frac{\phi'\Lambda\chi}{2}E_+' - (\delta - \frac{\phi'\chi}{2} - \frac{\Xi}{2})E_+)e^{-2i\phi}. \qquad (6.23)$$

The variation in the coefficients has thus been divided into two distinct parts. Firstly, there is the obvious rotational coefficient expressed as a complex exponential of twice the angle which the magnetic field makes with the x-axis, i.e. the factor on the extreme right of equations (6.22) and (6.23). Secondly, there is also position dependence contained in several quantities (Λ, χ, Ξ, κ and δ) because of ϕ', which is present in its own right but, for $\phi = \pi z/l$, ϕ' is constant.

8.3. Homogeneous magnetic field

For very large scalelengths, when $l \rightarrow \infty$, $\phi' \rightarrow 0$ and the odd order derivatives vanish, leading to the homogeneous dispersion relation as we will now demonstrate. When the gradient terms drop out, equations (6.22) and (6.23) reduce to:

$$E_+'' + \kappa E_+ = -\delta E_-(\cos 2\phi + i \sin 2\phi) \qquad (6.22a)$$

and

$$E_-'' + \kappa E_- = -\delta E_+(\cos 2\phi - i \sin 2\phi). \qquad (6.23a)$$

By using the definitions of E_+ and E_- given in equation (6.21), it may be readily established that E_x and E_y must satisfy:

$$E_x'' + (\kappa + \delta \cos 2\phi) E_x = -\delta \sin 2\phi \, E_y, \qquad (6.24)$$

$$E_y'' + (\kappa - \delta \cos 2\phi) E_y = -\delta \sin 2\phi \, E_x. \qquad (6.25)$$

Equation (6.24) may be rearranged to express E_y in terms of E_x and, since the coefficients are now all constant, the second derivative of the y coordinate of the perturbed electric field is simply given by:

$$E_y'' = \frac{1}{-\delta \sin 2\phi}(E_x^{iv} + (\kappa + \delta \cos 2\phi) E_x'').$$

It now only remains to eliminate the y-components by resubstitution into equation (6.25). The results of this look more familiar when written in terms of direction cosines and the ordinary and extraordinary mode notation, bearing in mind that $\tilde{R}\tilde{L}/\tilde{S}$ will reduce to RL/S since $\phi' = 0$, where we have:

$$\kappa + \delta \cos 2\phi \equiv P\mu^2 + \frac{RL}{S}\nu^2 \ , \ \kappa - \delta \cos 2\phi \equiv P\nu^2 + \frac{RL}{S}\mu^2 \ , \ \delta \sin 2\phi \equiv (P - \frac{RL}{S})\mu\nu,$$

to yield the fourth-order differential equation describing the uniform system:

$$E_x^{iv} + (P + \frac{RL}{S})E''_x + \frac{PRL}{S}E_x = 0.$$

This may be Fourier transformed to give the dispersion relation:

$$k_z^4 - (P + \frac{RL}{S})k_z^2 + \frac{PRL}{S} = 0,$$

which has solutions, $k_z^2 = P$ and $k_z^2 = \frac{RL}{S}$, which are the ordinary and extraordinary modes respectively, as we would expect.

9. The dispersion relation

The substitutions $E_+ = me^{i\phi}$ and $E_- = ne^{-i\phi}$ in equations (6.22) and (6.23) yield

$$m''(1-\frac{\Lambda\chi}{2}) + im'(2\phi'(1-\frac{\Lambda\chi}{2})-\chi) + m(\kappa + \tfrac{1}{2}(\phi'\chi+\Xi) - \phi'^2(1-\frac{\Lambda\chi}{2}))$$
$$= n''\frac{\Lambda\chi}{2} + n(\frac{\Lambda\chi}{2}\phi'^2 - \delta + \tfrac{1}{2}(\phi'\chi+\Xi)) \qquad (6.26)$$

and

$$n''(1-\frac{\Lambda\chi}{2}) - in'(2\phi'(1-\Lambda\frac{\chi}{2})-\chi) + n(\kappa + \tfrac{1}{2}(\phi'\chi+\Xi) - \phi'^2(1-\frac{\Lambda\chi}{2}))$$
$$= m''\frac{\Lambda\chi}{2} + m(\frac{\Lambda\chi}{2}\phi'^2 - \delta + \tfrac{1}{2}(\phi'\chi+\Xi)). \qquad (6.27)$$

Earlier, we required that our magnetic field rotate round the vertical axis according to $\phi = \pi z/l$. This judicious choice of phase variation for the "co-moving electric field components" has led to cancellation of the phase dependence from the coefficients of equations (6.26) and (6.27) - thereby removing the only remaining position dependence. We can now proceed to solve these equations in the usual way by performing a Fourier transform, since all the coefficients are constant, and thus our coupled differential equations become simultaneous algebraic equations. Taking $m \rightarrow ae^{i\sigma z}$, and $n \rightarrow be^{i\sigma z}$ produces two relations between the amplitudes of m and n (i.e. a/b , the actual values of a and b being determined by the auxiliary conditions of the particular problem):

$$a(-\sigma^2(1-\frac{\Lambda\chi}{2}) - \sigma(2\phi'(1-\frac{\Lambda\chi}{2})-\chi) + \kappa + \tfrac{1}{2}(\phi'\chi+\Xi) - \phi'^2(1-\frac{\Lambda\chi}{2}))$$
$$= b(-\sigma^2\frac{\Lambda\chi}{2} - \delta + \tfrac{1}{2}(\phi'\chi+\Xi) + \phi'^2\frac{\Lambda\chi}{2}),$$

and

$$b(-\sigma^2(1-\frac{\Lambda\chi}{2})+\sigma(2\phi'(1-\frac{\Lambda\chi}{2})-\chi)+\kappa+½(\phi'\chi+\Xi)-\phi'^2(1-\frac{\Lambda\chi}{2}))$$

$$= a(-\sigma^2\frac{\Lambda\chi}{2}-\delta+½(\phi'\chi+\Xi)+\phi'^2\frac{\Lambda\chi}{2}).$$

By cross-multiplication, we derive a quadratic in σ^2 which is our dispersion relation:

$$(1-\Lambda\chi)\sigma^4+(\,(\Lambda\chi-1)(\Xi+2\phi'^2)+\Lambda\chi(\kappa-\delta-\phi'\chi)-2\kappa+\chi(3\phi'-\chi)\,)\,\sigma^2+$$

$$(\kappa+\delta)(\kappa-\delta+\phi'\chi+\Xi+\Lambda\chi\phi'^2)+\phi'^4(1-\Lambda\chi)-\phi'^2(\phi'\chi+\Xi+2\kappa)=0. \quad (6.28)$$

This was one of the many points where the algebraic manipulation language, REDUCE (1987), proved to be an invaluable aid. The four eigenvalues, σ_j, which are the solutions of equation (6.28) thus provide the complete solution of our problem. Once they have been calculated, the ratio of the amplitudes may also be determined. In this way, we know m and n which are easily transformed into E_+ and E_-. These are easily rearranged into the true components of the electric field and these, in turn, may be related back to any of the other original variables as they are required (i.e. n_s, v_s, **B**, **J**).

In order to non-dimensionalise the dispersion relation and hence the eigenvalues, we now introduce a set of dimensionless variables which we will use in the remainder of this chapter. These include χ_0 and Ξ_0 which were encountered before:

$$\chi_0=\frac{\chi}{\phi'}\ ,\ \Lambda_0=\phi'\Lambda\ ,\ \Xi_0=\frac{\Xi}{\phi'^2}\ ,\ \kappa_0=\frac{\kappa}{\phi'^2}\ ,\ \delta_0=\frac{\delta}{\phi'^2}.$$

Similarly, a non-dimensional form of the eigenvalues may be obtained by defining $s=\sigma/\phi'$ and thus equation (6.28) becomes a quadratic in s^2 of the form:

$$\mathrm{a}s^4+\mathrm{b}s^2+\mathrm{c}=0, \quad (6.29)$$

where

$$\mathrm{a}=(1-\Lambda_0\chi_0),$$

$$\mathrm{b}=(\Lambda_0\chi_0-1)(\Xi_0+2)+\Lambda_0\chi_0(\kappa_0-\delta_0-\chi_0)-2\kappa_0+\chi_0(3-\chi_0),$$

$$\mathrm{c}=(\kappa_0+\delta_0)(\kappa_0-\delta_0+\chi_0+\Xi_0+\Lambda_0\chi_0)+1-\Lambda_0\chi_0-\chi_0-\Xi_0-2\kappa_0.$$

We may therefore express s^2 as:

$$s^2=\frac{-\mathrm{b}}{2\mathrm{a}}\pm\frac{\sqrt{\mathrm{b}^2-4\mathrm{ac}}}{2\mathrm{a}}.$$

As explained in §8.2, the solutions for E_x and E_y may be written in closed form as soon as the solutions for E_+ and E_- are known. Writing the four solutions of equation (6.29) as s_j ($j = 1..4$) and the four free parameters of the problem as a_j, it is clear that m may be expressed as $m = \sum_{j=1}^{4} a_j e^{is_j\phi}$. Also, n may be written $n = \sum_{j=1}^{4} c_j a_j e^{is_j\phi}$, where the c_j can be seen by inspection from one of the amplitude relations to be expressible as:

$$c_j = \frac{b_j}{a_j}$$

$$= \frac{(1-\frac{\Lambda_0\chi_0}{2})s^2+(2(1-\Lambda_0\frac{\chi_0}{2})-\chi_0)s-\kappa_0-\tfrac{1}{2}(\chi_0+\Xi_0)+1-\frac{\Lambda_0\chi_0}{2}}{\frac{\Lambda_0\chi_0}{2}s^2+\delta_0-\tfrac{1}{2}(\chi_0+\Xi_0)-\frac{\Lambda_0\chi_0}{2}}.$$

A general form for E_x can be generated in the following way:

$$\begin{aligned} E_x &= \tfrac{1}{2}(E_+ + E_-) \\ &= \tfrac{1}{2}\sum_{j=1}^{4} a_j e^{is_j\phi}(e^{i\phi}+c_j e^{-i\phi}) \\ &= \tfrac{1}{2}\sum_{j=1}^{4} a_j e^{i\sigma_j z}((1+c_j)\cos\phi + i(1-c_j)\sin\phi) \\ &\triangleq \sum_{j=1}^{4} E_{xj}, \end{aligned}$$

where we have reverted to the independent variable z in the exponential part of the solution to emphasise that our solution simply represents a travelling wave of the usual form (e^{ikz}) modulated by a periodic envelope. Thus the amplitude is not constant but is position-dependent, being controlled by the spatial rotation of the magnetic field.

An expression for E_y may be found in a similar manner to be:

$$E_y = \frac{1}{2i}\sum_{j=1}^{4} e^{i\sigma_j z}((1-c_j)\cos\phi + i(1+c_j)\sin\phi) \triangleq \sum_{j=1}^{4} E_{yj}.$$

This component of the electric field again demonstrates modulation by the rotating field, $\mathbf{B}_0$. The solutions have thus proven to be of the form predicted by analogy to Bloch's theorem.

In Figure (6.6) the real parts of the electric field components are plotted over a range of z. We calculate E_z from the expression originally used to eliminate it from the components of the wave equation, namely:

$$E_z = -\frac{1}{\varepsilon_{33}}(\frac{1}{\omega}(\tau_{31}E'_x + \tau_{32}E'_y) + \varepsilon_{31}E_x + \varepsilon_{32}E_y),$$

where the tensor elements were defined earlier. Differentiating the form of the solution for E_x gives:

$$E_x{}' = \sum_{j=1}^{4} E_{xj}{}'$$

$$= \sum_{j=1}^{4} i\,\sigma_j E_{xj} - \phi' E_y.$$

And similarly,

$$E_y{}' = \sum_{j=1}^{4} i\,\sigma_j E_{yj} + \phi' E_x.$$

Hence we may express E_z in terms of known quantities such as E_x and E_y:

$$E_z = \frac{i}{\tilde{S}}((\frac{\phi' \mathrm{T}_2}{\omega} - \tilde{D})(\mu E_y - \nu E_x) - \frac{\mathrm{T}_2}{\omega}\sum_{j=1}^{4} i\,\sigma_j(\mu E_{xj} + \nu E_{yj})).$$

The real part of E_z versus z is also plotted in Figure (6.6) and it may be seen that it is far from negligible. This situation is therefore very different from the one envisaged by Choudhury (1988) who chose to investigate modes with $E_z = 0$ in preference to others. In fact, the above expression for E_z indicates that it is unlikely ever to vanish.

10. Number of modes

In this section, we will work out how many real roots of the dispersion relation there are for different parameter values. Although the dispersion relation possesses four solutions, we will refer to its two modes since the remaining solutions simply describe propagation in the opposite sense. Unlike the acoustic-gravity waves of Chapter 4, where the gravitational potential provided an energy source which led to all waves having an exponential envelope, we are interested here only in purely real solutions to equation (6.29) corresponding to propagating, undamped modes. We are thus concerned with a subset of all possible solutions to equation (6.29) - the set of solutions which can pass through an infinitely long plasma without violating energy conservation.

There may be 0, 1 or 2 real and positive solutions of our problem depending on the signs and relative magnitudes of the component parts of equation (6.29). The important quantities may be seen from above to be:

$$C1 = -\frac{b}{2}a\ , \quad C2 = b^2-4ac\ , \quad C3 = (-\frac{b}{2a})^2 - \frac{(b^2-4ac)}{(2a)^2} = \frac{c}{a}.$$

The cases of most interest to us will be those where the theory for an inhomogeneous plasma produces markedly different results from the homogeneous one - cases where there is a discrepancy in the predicted number of propagating modes. In particular, we will look for an increase in the number of propagating modes from homogeneous to inhomogeneous theory since for such cases we contend that the dispersion relation for the homogeneous plasma predicts not only an incorrect number of modes but does not permit the existence of a root which is in fact allowed. In addition to the possibility of altering the number of roots, it must be borne in mind that our full expressions for the electric field components will always be modulated by a factor due to the changing orientation of the magnetic field and dependent on the scalelength. Thus, even when the expected number of roots is the same, the solution has a fundamentally different form.

10.1. Effect of changing the scalelength, l

Keeping n, B_0 and ϕ' fixed, we scan through a range of values for ω evaluating the wavenumbers of the ordinary and extraordinary modes (P and RL/S) and the equivalent inhomogeneous wavenumbers (σ_1^2 and σ_2^2). If we cannot find any value of ω for which there is a change in the number of roots, we repeat the procedure with a smaller scalelength, i.e. increased rotational effect. Changes in the number of roots may be subdivided into two classes - class A being an increase in the number of roots from the homogeneous to inhomogeneous case and class B being a decrease. The general trends which emerge from this examination are as we would expect. As the scalelength is decreased and the magnetic field completes a revolution in a shorter length of plasma, deviations from the homogeneous case become increasingly apparent. As l decreases, the number of class B contradictions increases. Also, as l is decreased further, class A changes can be observed too and begin to predominate. As a concrete example, with $B_0 = 0.1$T, $n = 10^{19}$m^{-3}, $l = 0.025$m and in the neighbourhood of the frequency where $\tilde{R}\tilde{L}/\tilde{S} = 0$ near ω_L, we find

1) for $\omega = 0.9542\times\omega_{pe}$ there is a class B contradiction; $P<0$, $RL/S>0$ but the squares of the solutions to the dispersion relation (6.29) are both negative. Where one root was predicted for a uniform plasma, we have no wavelike solutions in the inhomogeneous case.

2) for $\omega = 0.9607\times\omega_{pe}$ there is a class A contradiction; $P < 0$, $RL/S > 0$ and the squares of the solutions to the dispersion relation (6.29) are both positive. Instead of the single root of homogeneous theory, we expect to find two propagating modes when the magnetic field orientation is changing.

Because we must scan at finite frequency intervals, it is not possible to set exact limits on the scalelength above which no discrepancies will be observed. By continual adjustment of the frequency selection, we could continue to refine our estimate of this limit ad infinitum, but as an approximate guide, we will quote the results obtained for the same values of n, B_0 and frequency range as above. We will discuss the reasons for concentrating on this particular frequency range, near a zero of $\tilde{R}\tilde{L}/\tilde{S}$, shortly. Testing frequencies at intervals of $10^{-5}\omega_{pe}$ ($1.8\times10^7 s^{-1}$), the limits of the scalelength prove to be:

no changes (class A or B) for $l \gtrsim 0.8$m ,

no changes of class A for $l \gtrsim 0.029$m .

In theory, we may consider as small a scalelength as we wish since our analysis is not dependent on small deviations from uniformity i.e. slow variation. The limit calculated earlier for $|Z| \ll |Y|$ only concerns the validity of the approximations for the frequencies at which the functions χ etc had zeros. The first real limit which we will set on l is that we do not wish the equilibrium velocity of the plasma to become relativistic. Thus we require $v_{0e} < 0.1\times c$, say which imposes the restriction:

$$\frac{\phi' c\,\Omega_e}{\omega_{pe}^2} < 0.1,$$

which requires, for example that $l > 0.005$m for $B_0 = 0.1$T and $n = 10^{19}m^{-3}$.

10.2. Phase space curves

The three quantities C1, C2 and C3, mentioned above, contain sufficient information to determine the attributes of a point in solution space - dictating whether the point represents both modes propagating, only one or neither. We may therefore divide a plot of solution space into regions containing different numbers of roots by drawing the limiting curves defined by $C1 = 0$, $C2 = 0$ and $C3 = 0$. The exact forms of these curves are expressed below as functions of χ_0 and Λ_0 to be drawn in κ_0–δ_0 phase space and attention will be drawn to the parameter values which lead to singularities in any of the coefficients.

Special case - $\Lambda_0\chi_0 = 1$

The most obvious special case occurs when the coefficient of s^4 in equation (6.29) vanishes. In this case the eigenvalues are the solutions of:

$$s^2 = \frac{\Lambda_0\chi_0-1+\chi_0+\Xi_0+2\kappa_0-(\kappa_0+\delta_0)(\kappa_0-\delta_0+\chi_0+\Xi_0+\Lambda_0\chi_0)}{(\Lambda_0\chi_0-1)(\Xi_0+2)+\Lambda_0\chi_0(\kappa_0-\delta_0-\chi_0)-2\kappa_0+\chi_0(3-\chi_0)},$$

so that there can be at most one mode of propagation and there will be *no* undamped mode if the above expression is negative, i.e. if c and b of equation (6.29) have the same sign.

C1 = 0

This is a straight line defined by

$$\delta_0 = \kappa_0\left(1-\frac{2}{\Lambda_0\chi_0}\right)+2+\Xi_0-\chi_0-\frac{((\chi_0-1)(\chi_0-2)+\Xi_0)}{\Lambda_0\chi_0}.$$

This formula applies whenever both Λ_0, $\chi_0 \neq 0$. Recalling the definition of $\Lambda = v_{0e}/\Omega_e$, we recognise that Λ_0 will be zero if and only if $\phi' = 0$. In other words, $\Lambda_0\chi_0 = 0$ for a rotating field only if $\chi_0 = 0$. In this case ($\chi_0 = 0$, $\Lambda_0 =$ anything), the line C1 = 0 becomes the vertical line $\kappa_0 = -1-\Xi_0/2$.

C2 = 0

This is a parabola defined by:

$$d_2\delta_0^2+d_1\delta_0+d_0 = 0,$$

where

$$d_2 = (\Lambda_0\chi_0-2)^2,$$

$$d_1 = 2\chi_0(\Lambda_0\chi_0(\Lambda_0\chi_0+\chi_0-\Lambda_0(\kappa_0+\Xi_0)-1)+\Lambda_0(2\kappa_0+3\Xi_0)-2)-4\Xi_0,$$

$$d_0 = \Lambda_0^2\chi_0^2\kappa_0^2+2\kappa_0(\Lambda_0\chi_0(\Lambda_0\chi_0(4+\Xi_0)-\chi_0^2(1+\Lambda_0)+7\chi_0-\Xi_0-12)+2(\chi_0-2)^2)+$$
$$\Xi_0(\Xi_0(\Lambda_0\chi_0-1)^2+4(\Lambda_0\chi_0-1)(\Lambda_0\chi_0-2)+2(\chi_0-1)(\chi_0-2)+2\Lambda_0\chi_0^2(4-\Lambda_0\chi_0-\chi_0)-4)+$$
$$\chi_0^2(1+\Lambda_0)(\chi_0^2(1+\Lambda_0)-2\chi_0(2\Lambda_0+3)+12)+\chi_0(\chi_0-8).$$

If the coefficient of δ_0^2 in this equation is zero (i.e. $\Lambda_0\chi_0 = 2$), then the curve C2 = 0 retains its parabolic shape but has the vertical axis as an axis of symmetry since we may now write:

$$\delta_0 = -\frac{4\kappa_0(\kappa_0+\chi_0+\Xi_0)+\Xi_0(\Xi_0-2\chi_0(\chi_0-1))+\chi_0((\chi_0+2)(\chi_0-2)^2+\chi_0-8)}{4\chi_0^2}.$$

C3 = 0

Having already considered the special case of $\Lambda_0\chi_0 = 1$, we may therefore assume that the denominator of C3 is always non-zero and so C3 = 0 at the points where $c = 0$. The curve C3 = 0 is then given by :

$$\delta_0^2 - \delta_0(\chi_0(1+\Lambda_0)+\Xi_0) - \kappa_0^2 + \kappa_0(2-\chi_0(1+\Lambda_0)-\Xi_0) + \chi_0(1+\Lambda_0) + \Xi_0 - 1 = 0.$$

This does not represent a parabola as might initially be expected since this quadratic factors easily making C3 = 0 true on a pair of straight lines with the equations:

$$\delta_0 = \kappa_0 + \chi_0(1+\Lambda_0)+\Xi_0 - 1 \quad \text{and} \quad \delta_0 = -\kappa_0 + 1.$$

Finally, we note that the points of intersection of the lines are common to all three lines because:

$$\mathrm{C1} = 0 \Rightarrow \mathrm{b} = 0 \qquad \mathrm{C3} = 0 \Rightarrow \mathrm{c} = 0$$

and

$$\mathrm{b} = 0\,,\;\; \mathrm{c} = 0 \Rightarrow \mathrm{C2} = 0.$$

10.3. Number of modes

Depending on which of the three quantities C1, C2, C3 are positive and which are negative, we may state the number of positive, real roots of equation (6.29). Expressing equation (6.29) as:

$$s^2 = \mathrm{C1} \pm \frac{\sqrt{\mathrm{C2}}}{2\mathrm{a}}, \tag{6.30}$$

we see that if C2 < 0 there cannot be a purely real eigenvalue and we say that there are no propagating modes. In order to determine which of the two terms on the right hand side of equation (6.30) is larger, we calculate:

$$\mathrm{C1}^2 - \frac{\mathrm{C2}}{4\mathrm{a}^2} = \frac{\mathrm{c}}{\mathrm{a}} = \mathrm{C3}.$$

For instance, if all three of C1, C2 and C3 are positive then both modes of propagation are possible. If C3 alone were negative, the discriminant would be larger than the C1 term, forcing the "minus" root to satisfy $s^2 < 0$ and so there would be a single allowed mode. A complete list of the various possibilities is most easily demonstrated by the table below (where a "+" sign indicates a quantity greater than zero and "-" indicates a negative one) :-

C_1	C_2	C_3	No. of Modes
+	+	+	2
+	+	-	1
-	+	-	1
-	+	+	0
±	-	±	0

Figure (6.7) shows a series of plots in κ_0–δ_0 phase space with the limiting curves given by C1 = 0, C2 = 0 and C3 = 0 as shown, separating the regions containing 0, 1 or 2 real roots which are indicated by the labels 0, 1 and 2, respectively.

This diagram is parametrised by Λ_0, χ_0 and Ξ_0. The four free variables of the problem are n, B_0, ω and ϕ'. Using two of these to change the axes variables leaves us free to set two of Λ_0, χ_0 and Ξ_0 and thus the final one is prescribed. Our choice of these variables will be governed by the values which arose naturally from the numerical calculations used to study the effect of changing the scalelength.

Figure (6.7a) shows the division of κ_0–δ_0 space for $\phi' = 0$, where the homogeneous dispersion relation predicts that there will be two real roots for $P > 0$ and $RL/S > 0$, one mode when only one is positive and no roots when both P and RL/S are negative. (Note that to translate this graph to the more familiar P, RL/S space simply requires a rotation of 45^o about the origin.)

In Figure (6.7b), $\Lambda_0 = 1.0\times10^{-4}$, $\chi_0 = 0.0$ and $\Xi_0 = 1.0\times10^{-2}$ which represents a reasonably long scalelength (l = 0.1m for n = 10^{19}m^{-3}) and therefore a relatively small change from homogeneity. It can be seen that changes in the number of modes are most likely to occur along the main diagonals, $\kappa_0 = \pm\delta_0$, corresponding to $P = 0$ and $\tilde{R}\tilde{L}/\tilde{S} = 0$. This is why we selected $\omega \approx \omega_L$ in the earlier examples since, for the parameters chosen, $\tilde{R}\tilde{L}/\tilde{S}$ has a zero very near ω_L. In fact, the zero of P lies at ~$1.00027\times\omega_{pe}$, $\omega_L \approx 0.95\omega_{pe}$ and $\omega_R \approx 1.05\omega_{pe}$ so that the neighbourhood of the electron plasma frequency, containing these three zeros of P and $\tilde{R}\tilde{L}/\tilde{S}$, is likely to contain the bulk of the changes from the homogeneous case. Indeed, for these parameters, the regions where contradictions occur for comparatively large values of l are all found at the frequencies of the 5 zeros of P and $\tilde{R}\tilde{L}/\tilde{S}$ (frequencies which were identified in § 7.1.3).

In the following two plots, we consider the locations of two specific points in phase space and check that the number of modes which they indicate correspond to the predicted number of solutions calculated from equation (6.29). (For clarity, we force the inner regions of phase space to be of comparable sizes by avoiding very small values of the product $\Lambda_0\chi_0$, since this term occurs in the denominator of the curve C1 = 0. This necessitates using extremely small scalelengths but these examples are intended purely to be illustrative of the overall behaviour.)

Figure (6.7c) portrays $\Lambda_0 = 2.79$, $\chi_0 = 0.26$ and $\Xi_0 = 0.42$. Let us consider the point $\kappa_0 = -2.8\times10^{-2}$, $\delta_0 = 2.9\times10^{-2}$. According to the graph, this point corresponds to a single propagating mode for the inhomogeneous plasma. However, since this point satisfies $P > 0$, $RL/S > 0$, two modes could propagate under these conditions in a uniform plasma. That this is indeed a contradiction of type B may be checked by substituting these values of κ_0 and δ_0 into the definitions of C1, C2 and C3 which results in C3<0 and the others positive. Clearly from the table, only one mode may propagate. Alternatively, we may calculate the values directly, yielding:

$$P \approx 79 \quad \frac{RL}{S} \approx 3.6\times10^4 \quad \sigma_1^2 \approx -6.9\times10^3 \quad \sigma_2^2 \approx 9.1\times10^3,$$

which agrees with the other predictions.

Figure (6.7d) shows a similar situation but in this case we have chosen to highlight an *increase* in the number of roots from the homogeneous case. Here, $\Lambda_0 = 2.79$, $\chi_0 = 0.36$ and $\Xi_0 = 0.49$ and we are concerned with the point $\kappa_0 = 3.5\times10^{-2}$, $\delta_0 = 1.05\times10^{-3}$ which lies in the section for two modes. Since $P > 0$ ($P \approx 3.6\times10^3$) but $RL/S > 0$ ($RL/S \approx -8.6\times10^5$), the equivalent parameters would only permit one mode in a homogeneous plasma. Again, we check that both solutions of equation (6.29) are positive:

$$\sigma_1^2 \approx 801 \qquad \sigma_2^2 \approx 31.$$

The remaining parts of Figure (6.7) show how altering the values of Λ_0, χ_0 and Ξ_0 affect the topology of these diagrams. Their values have been assigned arbitrarily in order to make the figures as clear as possible and have not been derived from specific choices of the physical parameters.

11. Summary

In this chapter we have made a natural progression from the discussion of Chapter 5 by considering the effect of inhomogeneity, through a "rotating" magnetic field, on the natural modes of perpendicular propagation in a cold plasma. The first effect on which we remarked was the introduction of a non-zero plasma current, which was required to balance the current due to the sheared magnetic field. We assigned this current to the motion of both ions and electrons by requiring that the nett momentum of the system remain zero - thus the bulk of the current is due to electrons moving at $-m_i/m_e$ times the ion velocity.

This equilibrium current, in turn, coupled to the other system variables necessitating the inclusion of density perturbation effects and resulting in a more complicated Ohm's Law than for the uniform case. The resultant change in the dielectric (and other) tensor elements was accompanied by a frequency shift in the denominator of these terms - again an effect proportional to the equilibrium velocity. The equations for determining the electric field components were found to form a coupled pair of second order ode's, but the periodic nature of the coefficients made solution by standard methods impossible. It was at this stage that a transformation of the electric field components into an analogous pair, mimicing the rotation of the ambient magnetic field, was performed and this provided the key to the solution of the problem. These constitute a more natural set of dependent variables for the system and solution from this stage was straightforward. The one restriction was the requirement that ϕ' be constant.

The solutions of the equations appear to comply with all our expectations - their general nature being propagating waves modulated by a periodic envelope. This result agrees with the predictions extrapolated from Bloch's theorem which applies to a periodic crystal lattice. As the scalelength of rotation decreases, the likelihood of a different number of real solutions arising in the inhomogeneous case than would be predicted by the homogeneous dispersion relation grows. This was demonstrated graphically and by numerical examples.

The waves which propagate in such a rotating magnetic field are obviously very different from the ordinary and extraordinary modes. Our analysis has demonstrated the danger of oversimplification in the study of nonuniform plasmas and of relying solely on extensions of conclusions derived in the homogeneous environment.

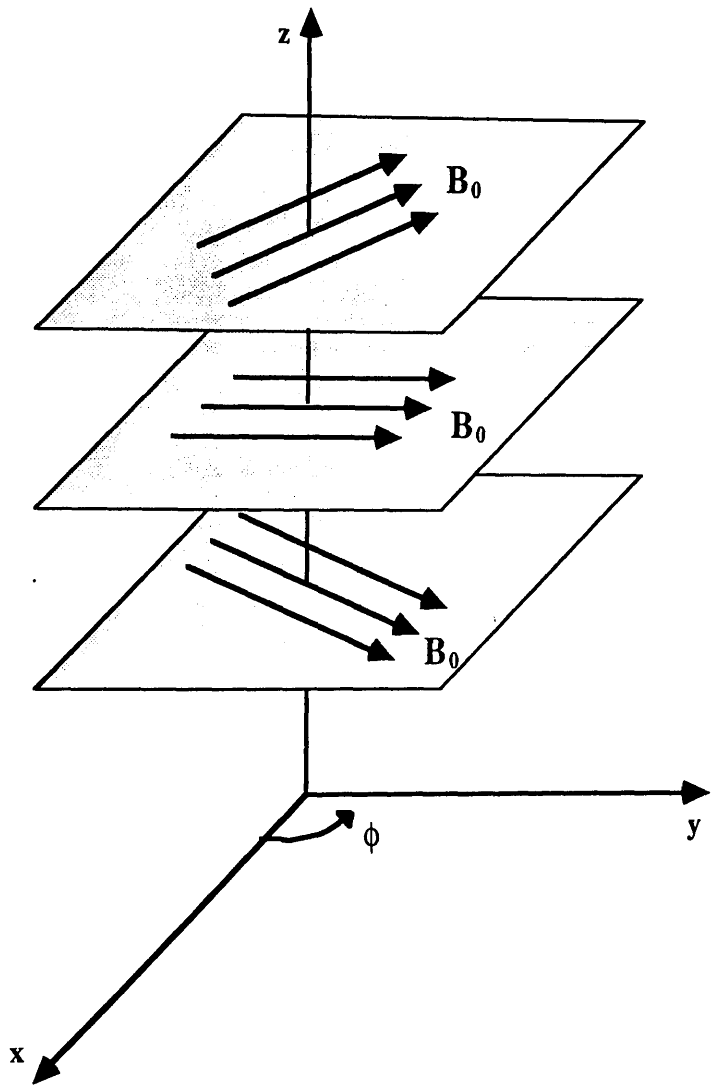


Figure (6.1) Variation of the "spatially rotating" magnetic field, showing its different orientation at three selected positions and its constant magnitude.

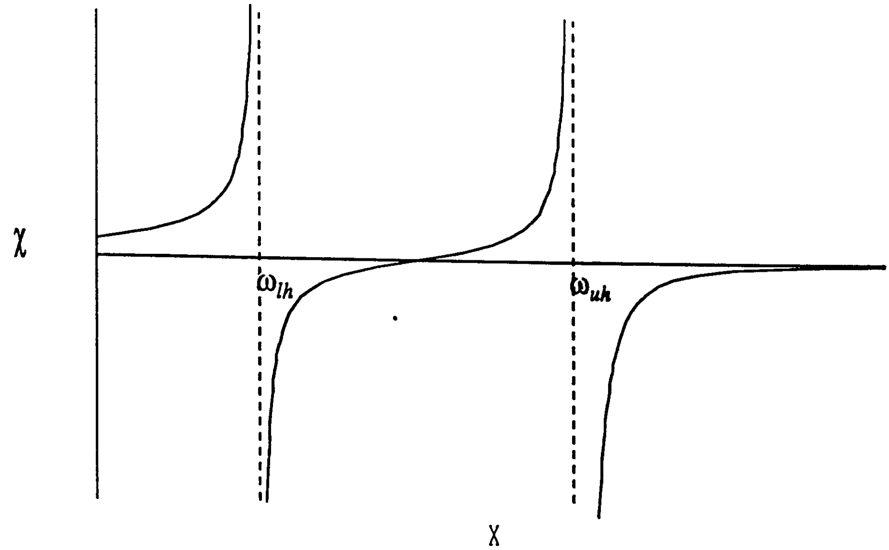


Figure (6.2) Schematic diagram of the variation of the quantity, χ, with dimensionless frequency, $X = \dfrac{\omega}{\omega_{pe}}$.

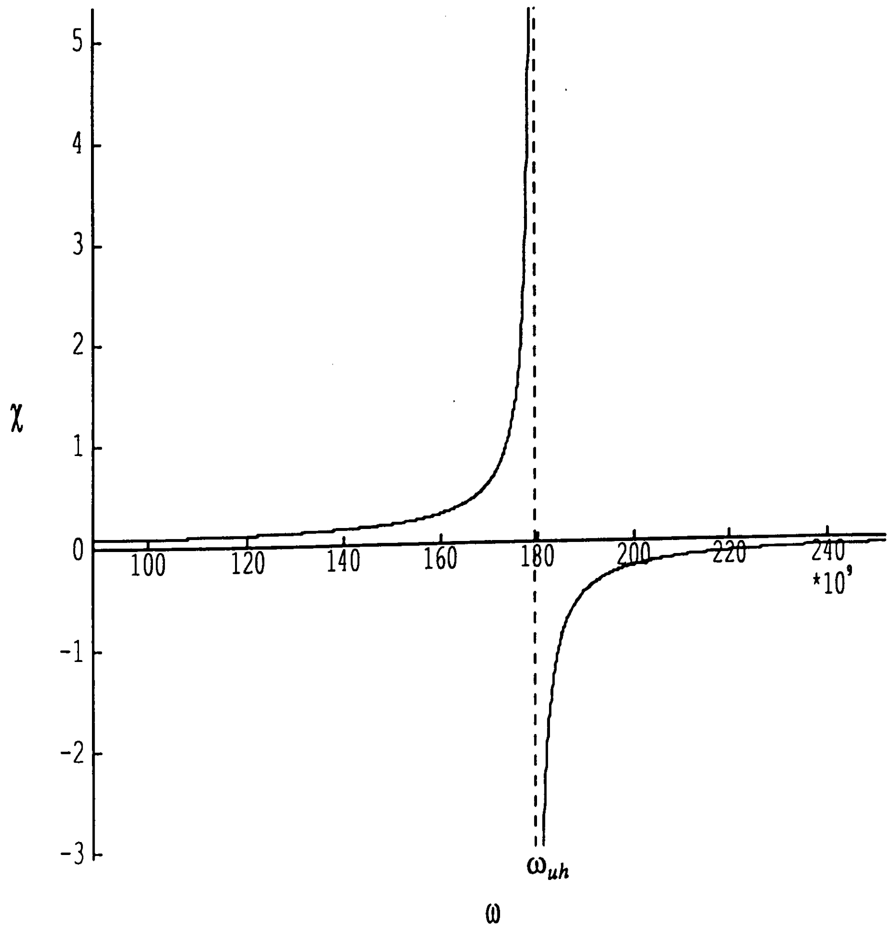


Figure (6.3) Plot of the behaviour of χ in the vicinity of the upper hybrid resonance, illustrating one of the resonances of χ. (Parameters are $B_0 = 0.1$T, $n = 10^{19}\mathrm{m}^{-3}$, for which $\omega_{uh} = 1.7926\times10^{11}\mathrm{s}^{-1}$.)

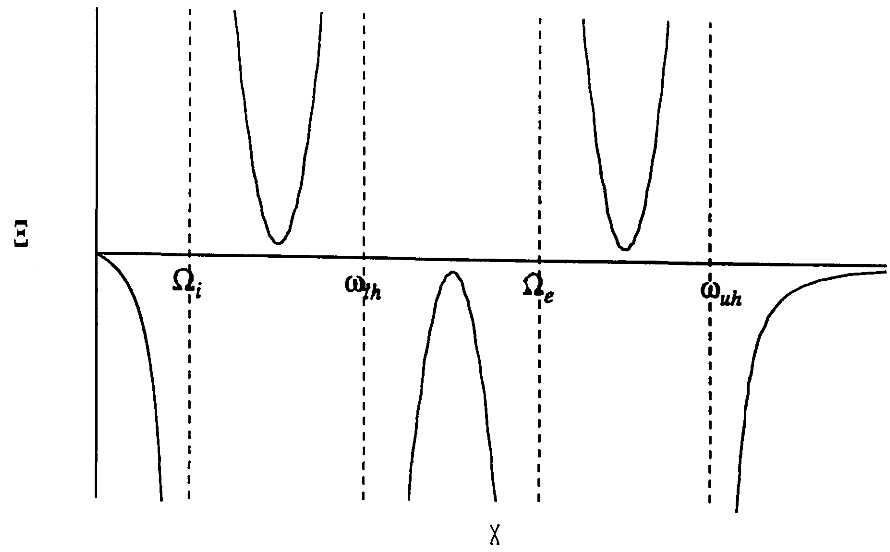


Figure (6.4) Schematic diagram of the variation of the quantity, Ξ, with dimensionless frequency showing that it has 4 resonances and no zeros.

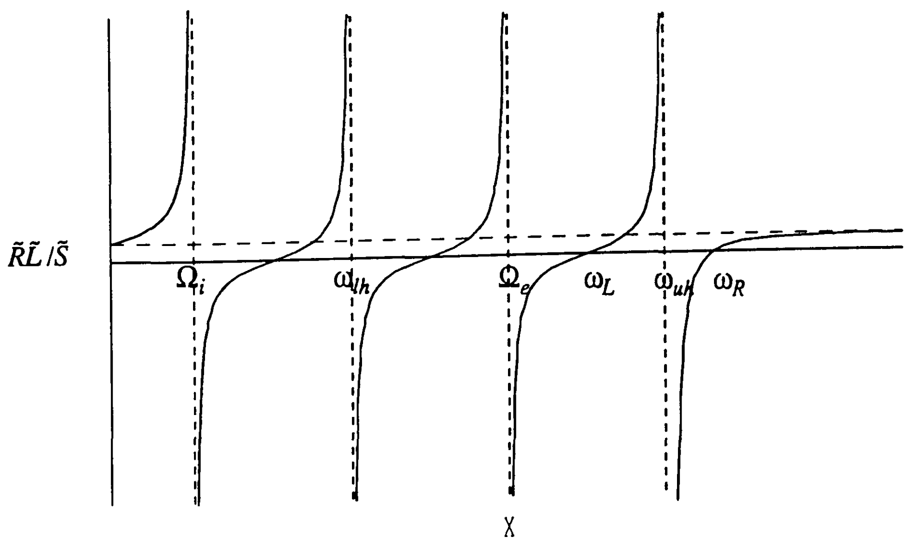


Figure (6.5) Schematic diagram of the variation of the quantity, $\tilde{R}\tilde{L}/\tilde{S}$, with dimensionless frequency. (Contrast this with Figure (5.3) showing RL/S versus X).

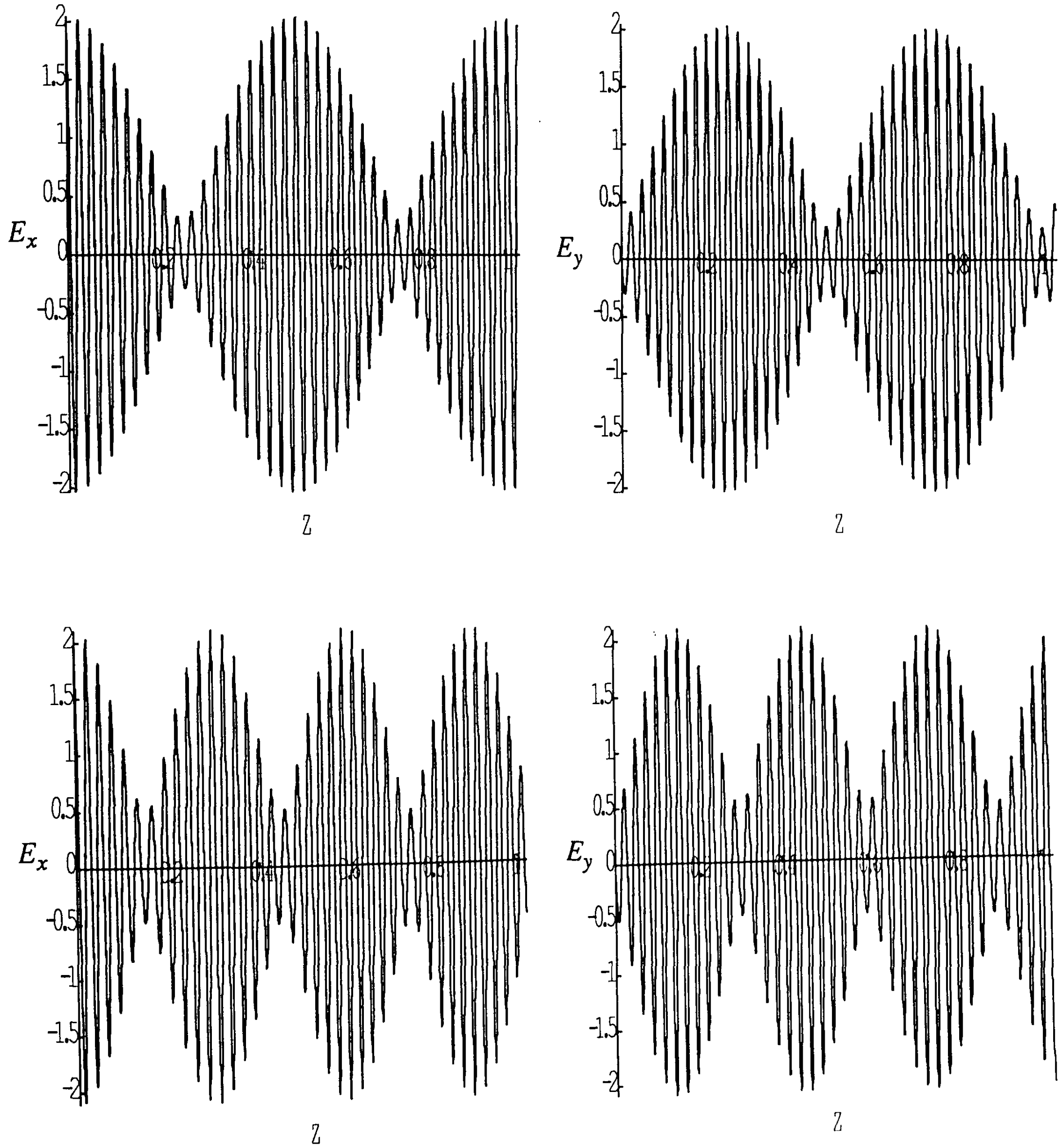


Figure (6.6a) Variation of the x and y components of the electric field with position in a cold plasma with a spatially rotating magnetic field. The top pair have scalelength $l = 0.5$m and the bottom pair have $l = 0.3$m so that the length of the envelope changes between the two pairs. The symmetry of these graphs results from the choice of amplitudes, $a_1 = a_3 = 1$ and $a_2 = a_4 = 0$. Note the phase difference between E_x and E_y.

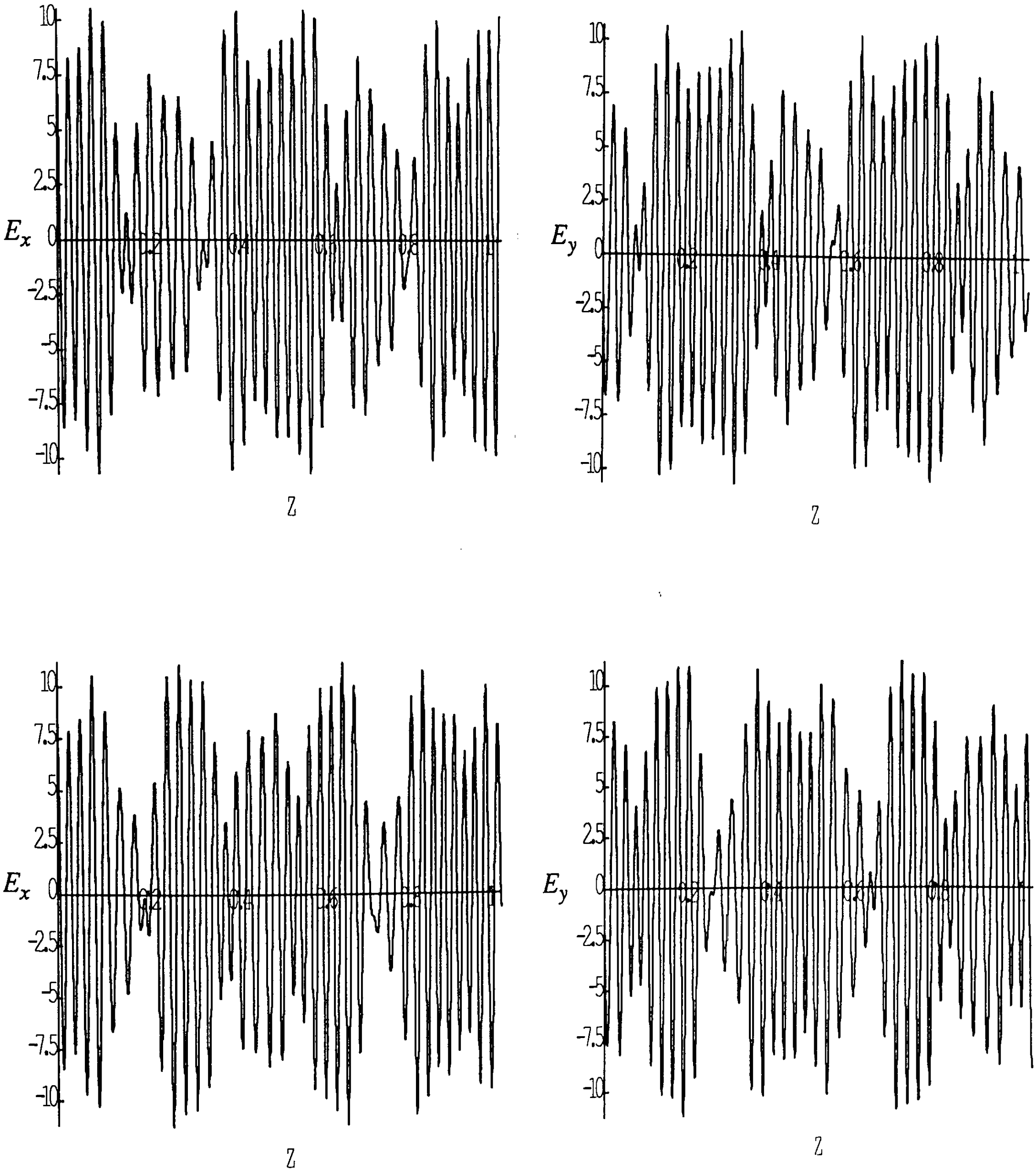


Figure (6.6b) Variation of the x and y components of the electric field with position in a cold plasma with a spatially rotating magnetic field. The top pair have scalelength $l = 0.5$m and the bottom pair have $l = 0.3$m so that the length of the envelope changes between the two pairs. The amplitudes chosen, $a_1 = 5$, $a_2 = -2$, $a_3 = 4$ and $a_4 = 3$, lead to more complicated structure than the symmetric amplitudes of Figure (6.6a).

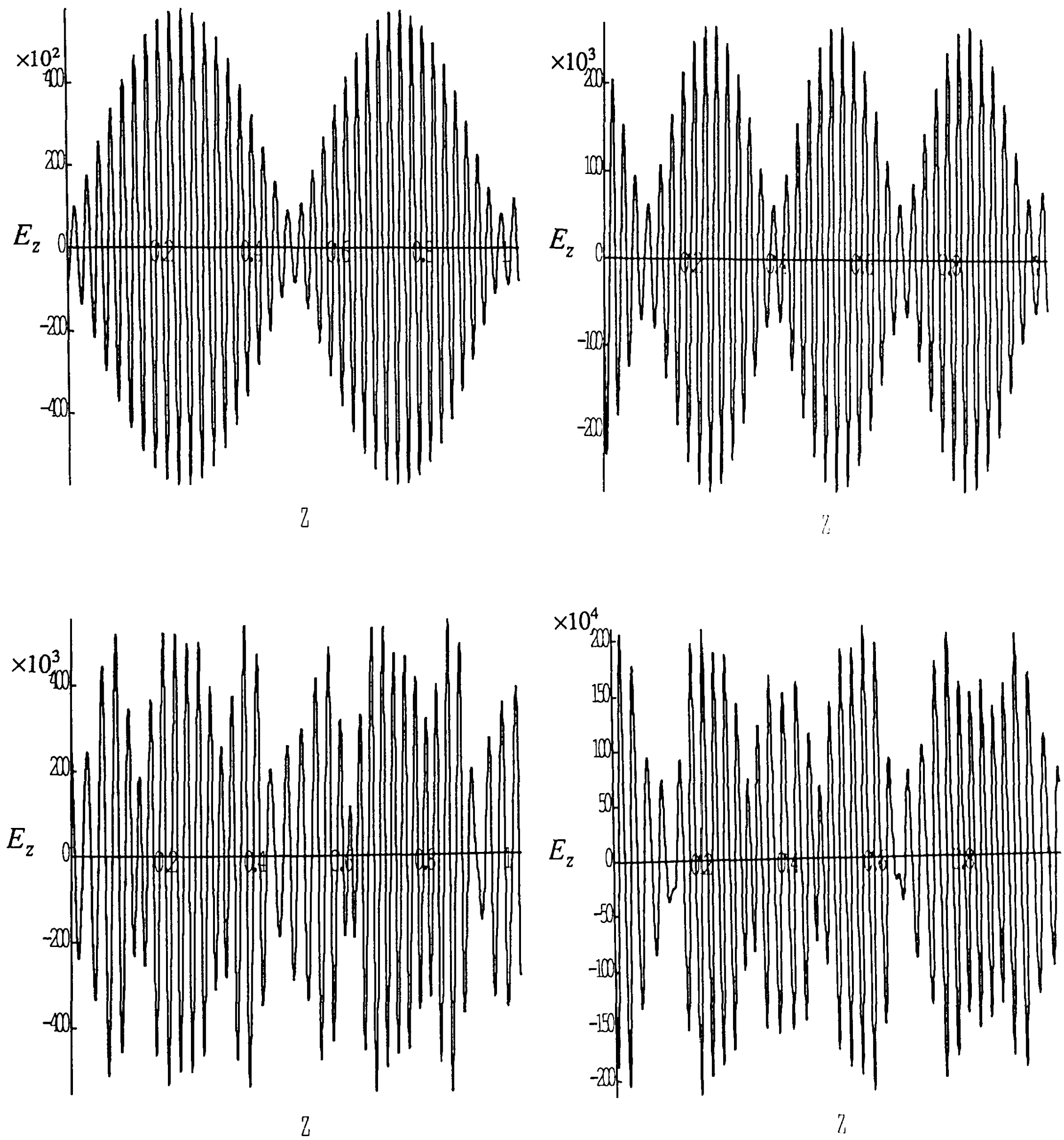


Figure (6.6c) Variation of the z component of the electric field with position in a cold plasma with a spatially rotating magnetic field. The two graphs on the left depict l = 0.5m and those on the right, l = 0.3m. The top pair have the same amplitudes as Figure (6.6a) and those below the amplitudes of Figure (6.6b).

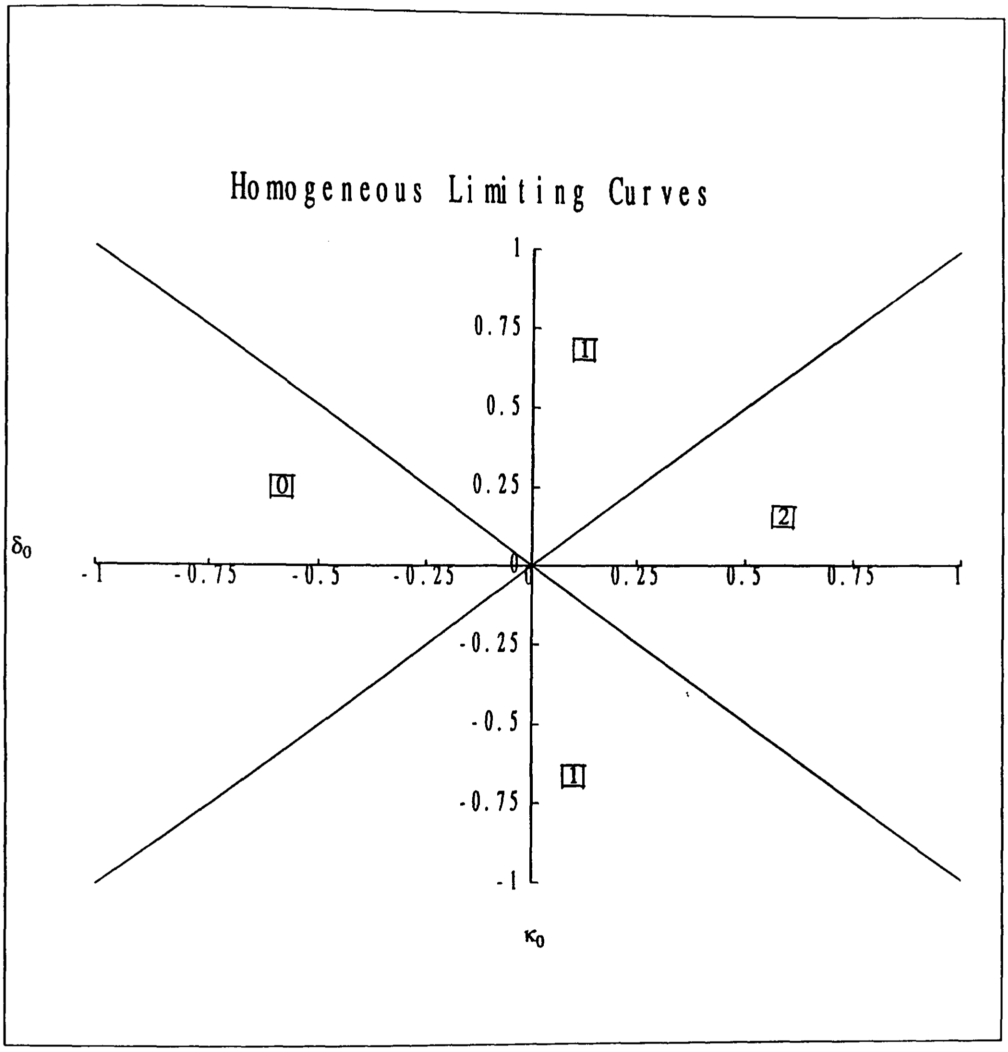


Figure (6.7a) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region. This diagram represents homogeneous phase space where the magnetic field is constant.

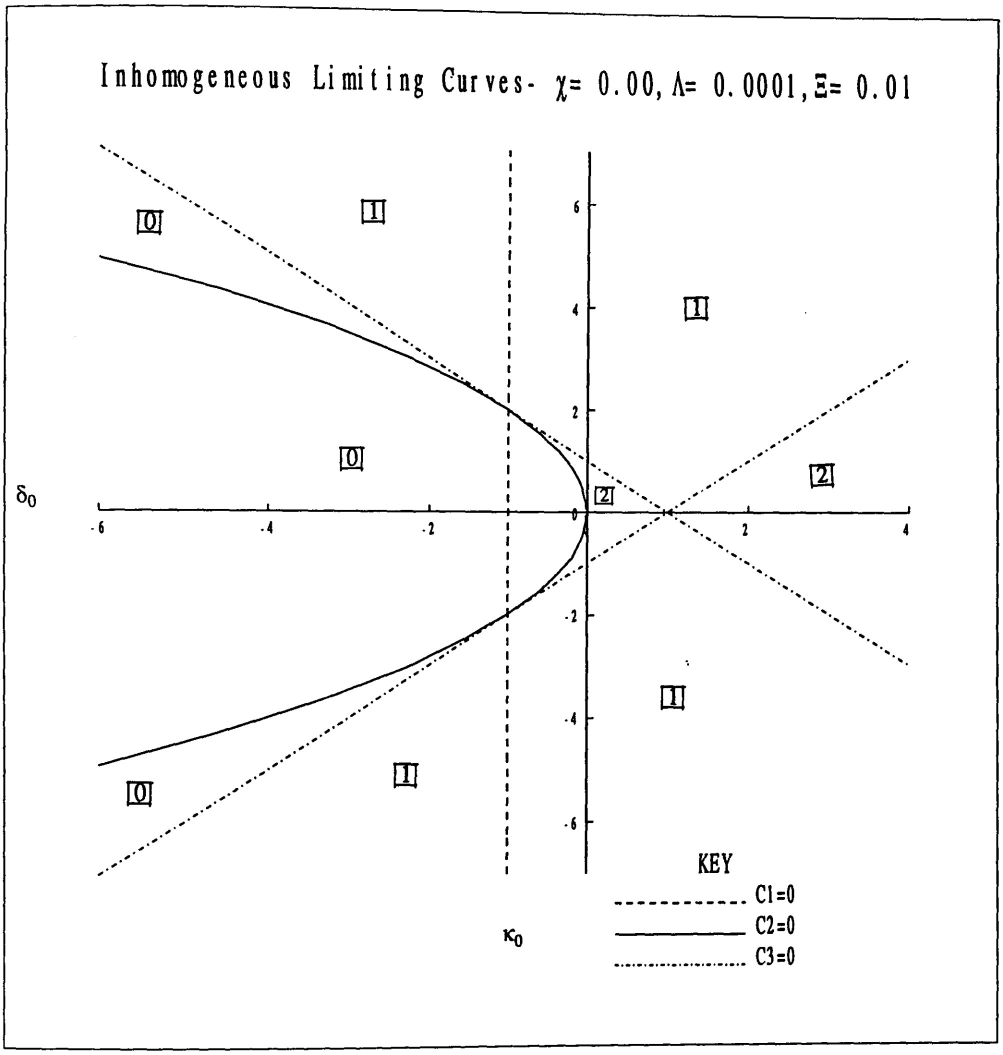


Figure (6.7b) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region. This diagram represents a small deviation from homogeneity.

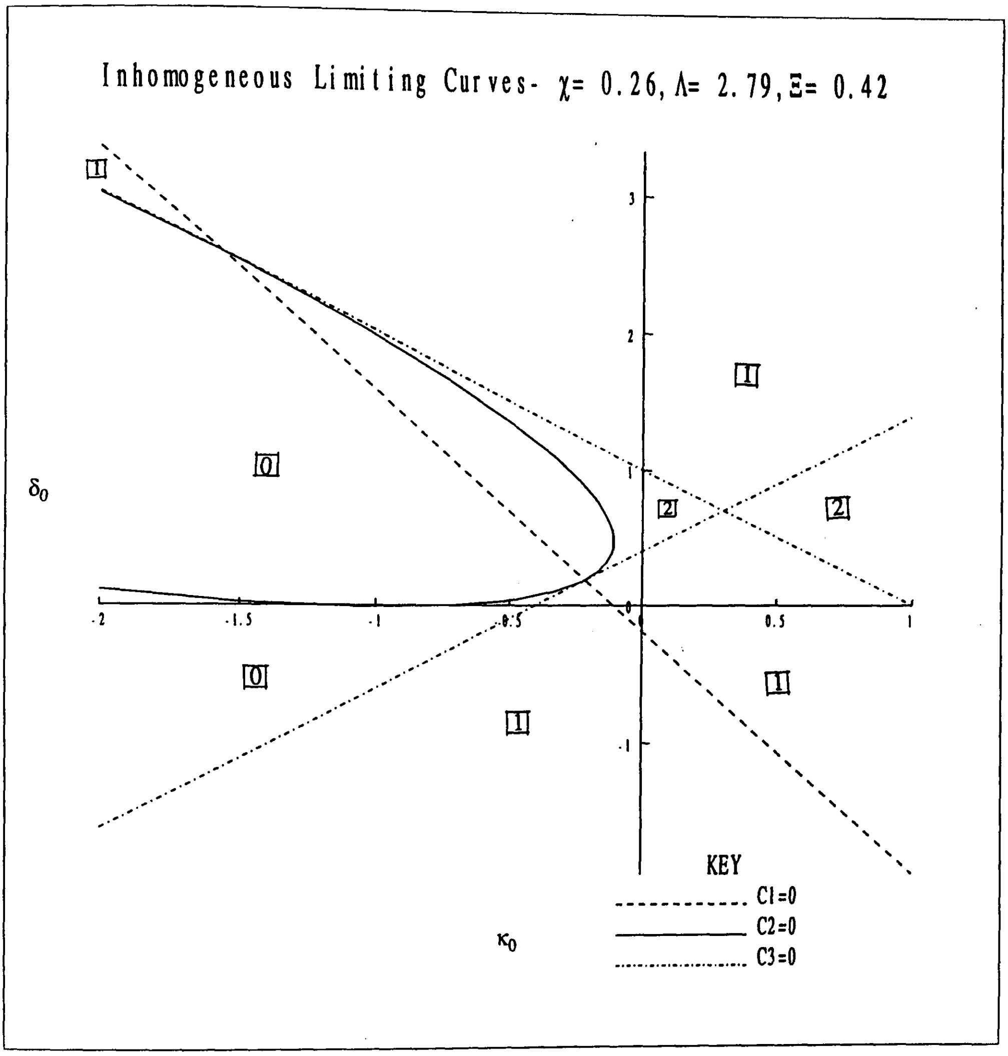


Figure (6.7c) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region.

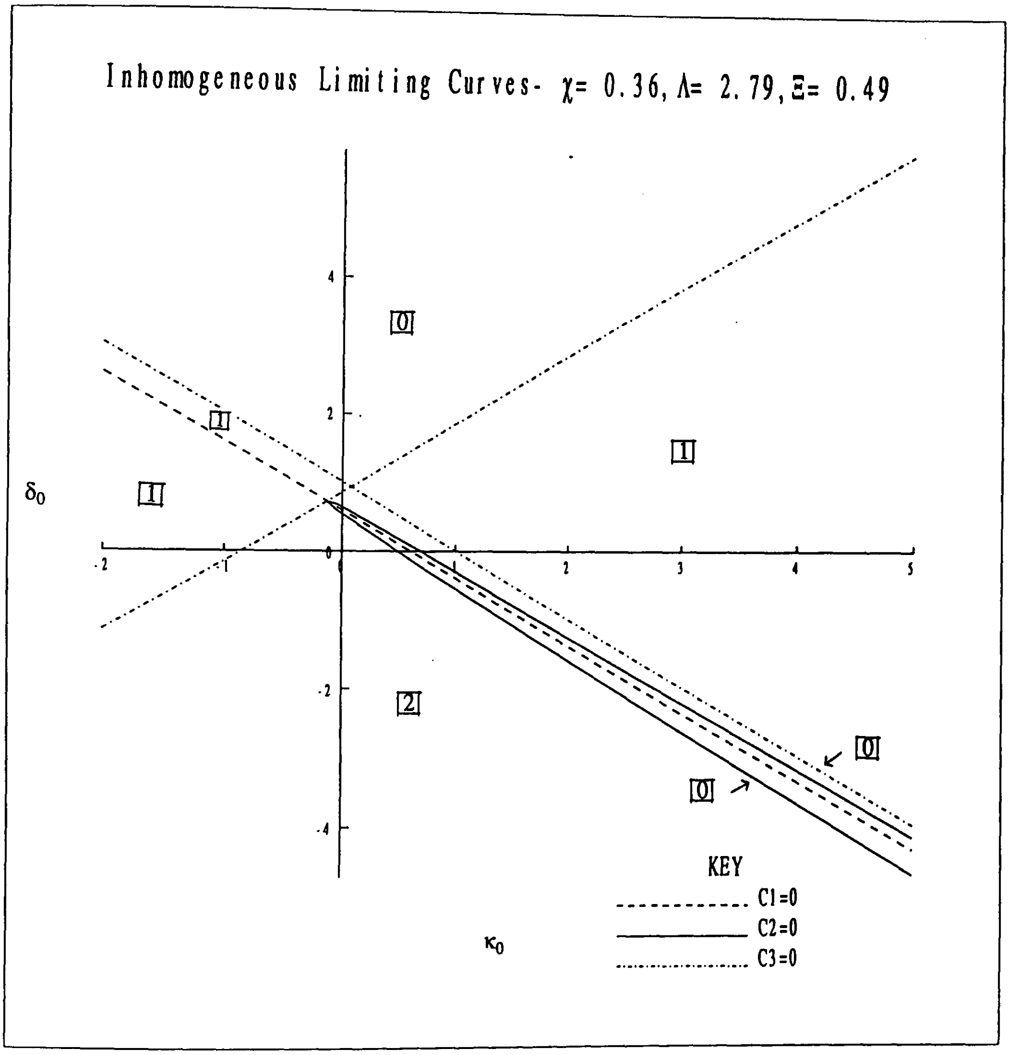


Figure (6.7d) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region.

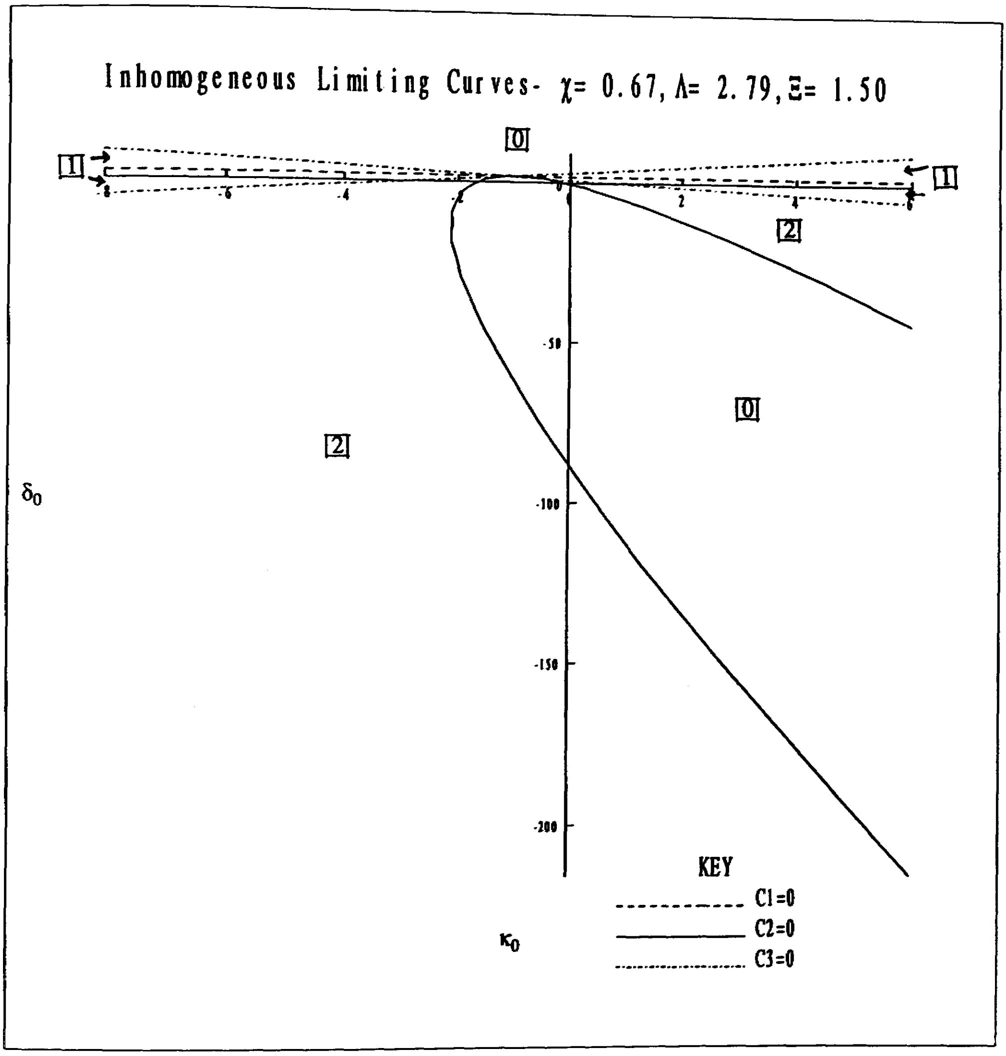


Figure (6.7e) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region. The parameters for this diagram were chosen so that $\Lambda\chi \sim 2$, which is a limiting value beyond which the parabola C2 will be open to the left.

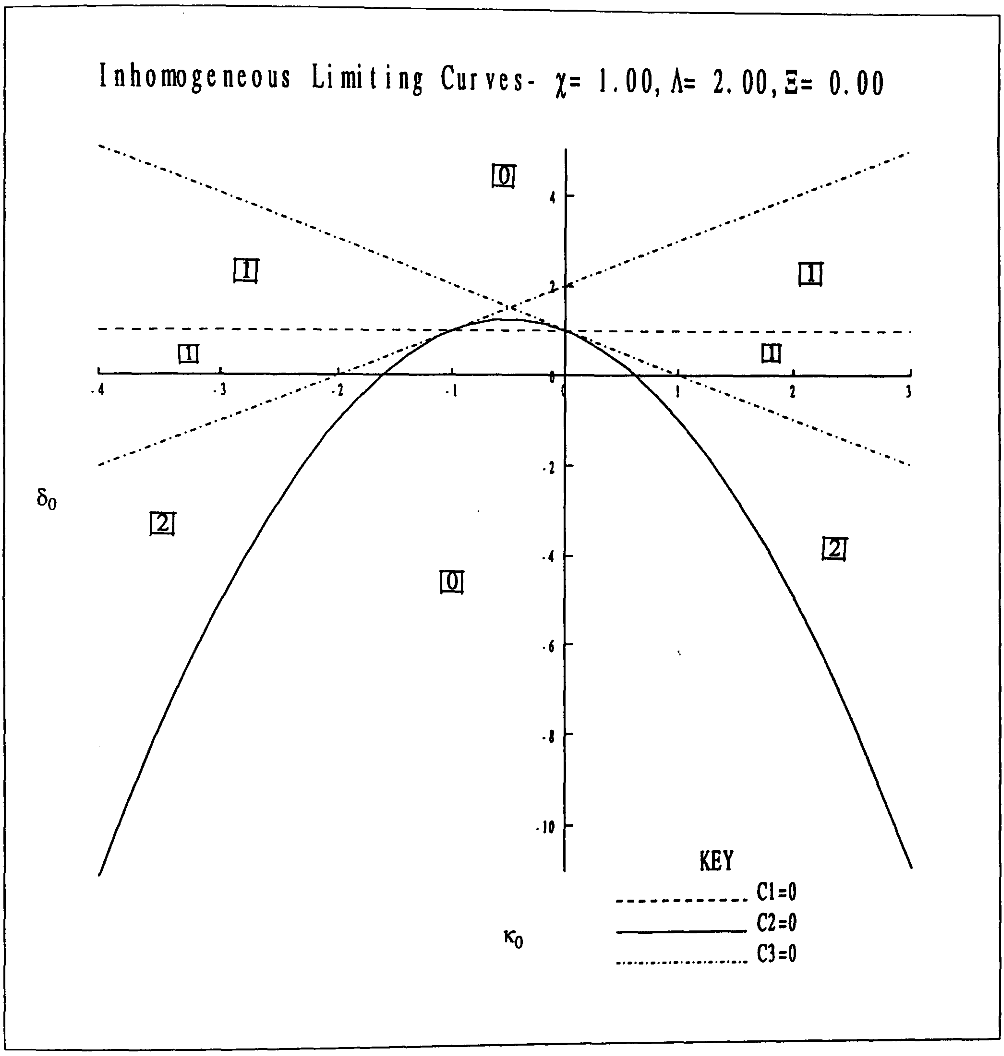


Figure (6.7f) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region. For $\Lambda\chi = 2$, the parabola C2 is symmetric about the δ_0 axis.

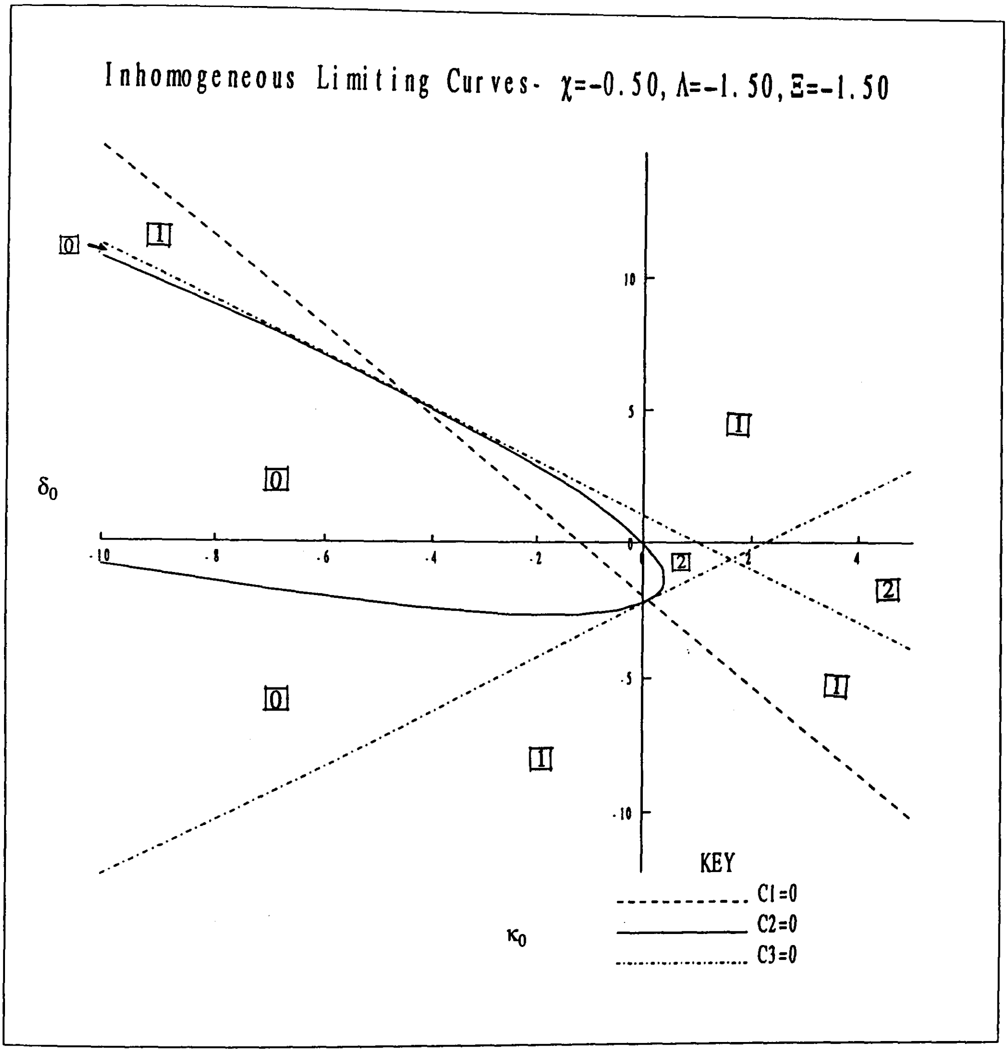


Figure (6.7g) Division of κ_0–δ_0 space by the lines C1 = 0, C2 = 0 and C3 = 0 showing the number of modes possible in each region.

Chapter 7 - Future Work

1. Possible future work

Having examined in detail the change in propagation characteristics of waves in a cold plasma with a spatially rotating magnetic field, we will now consider briefly how this analysis may be utilised and two ways in which it may be extended. We will indicate how these extensions may be included within the model but will leave the solution of the problems they pose to future investigations.

In Chapter 2 we discussed mode conversion and the need for a method which is better than the inverse Fourier transform technique for studying wave propagation in media with position-dependent characteristics. In this work we have derived the exact solution to the problem of waves travelling in a magnetised plasma where the propagation is perpendicular to the plane of the magnetic field but the orientation of the ambient magnetic field varies. Given the details of the field - its strength and variation with position - plus the plasma density, we can predict the behaviour of a wave of a given frequency and the values of the field variables at any point.

2. Application to mode conversion

One problem of great interest still remains - how much mode conversion would result from passing a wave through a plasma with the magnetic field described in Chapter 6? Because the system is of fourth order and therefore can support two entirely different wavemodes, the result of any mode conversion could be a wave of completely different form from the input wave instead of simply a reflection of the incident wave - the result of mode conversion in a second order system. The eigenvalues are known throughout the plasma, whether it be uniform or inhomogeneous, since they are defined by the plasma and magnetic field parameters at each point. Thus, in theory, any wave propagation problem in this context may be solved simply by calculating the steady state solution and from this determining how much, if any, of a secondary mode was excited by the passage of the incident wave through the inhomogeneity.

Because this is a true mode conversion problem with four solutions, representing both forward and backward propagating modes, it is highly complicated and is beyond the scope of the present thesis. Currently, work is being carried out in this area by Diver and Laing (1989), who are observing the change in a wave of initially prescribed wavenumber emerging from uniform plasma into a section of inhomogeneity and finally returning to a spatially-constant region. Initial results indicate that deriving the coefficients of reflection and transmission will require a great deal of effort and, although theoretically straightforward, this process will undoubtedly require considerable use of computer algebra in the extensive manipulations of the variables which is involved. Matching the derivatives (up to third order) of the variables at the boundaries between the homogeneous and inhomogeneous regions provides sufficient auxiliary conditions to complete the specification of the problem. In an analogous way to quantum mechanical scattering, the reflection and transmission coefficients may then be calculated, although it must be emphasised that this is a far from trivial extension of "potential step" problems.

3. Changing rate of rotation - $\phi'' \neq 0$

One natural method of extending the work of Chapter 6 would be to remove the restriction limiting the ambient magnetic field variation to a linear relationship with z. Thus ϕ'' would no longer be zero, causing repercussions throughout the analysis and introducing additional terms at the earliest stages. We now repeat the fundamental steps of the analysis of Chapter 6, assuming - $\phi'' \neq 0$, to illustrate the influence of these new terms.

The continuity and momentum equations are unchanged from Chapter 6 (as are Maxwell's equations) so that we may again eliminate the velocity in favour of the electric field and its first derivative:

$$\mathbf{v}_s = \frac{q_s}{m_s}\mathbf{M}^{-1}\mathbf{E} + \frac{q_s}{m_s}\mathbf{M}^{-1}\cdot\mathbf{N}\,\mathbf{E}', \tag{6.13}$$

where $\mathbf{M}$ and $\mathbf{N}$ are defined as before to be:

$$\mathbf{M} = \begin{bmatrix} -i\omega & 0 & \varepsilon_s\Omega_{sy} + \mathrm{v}_{0sx}' \\ 0 & -i\omega & -\varepsilon_s\Omega_{sy} + \mathrm{v}_{0sy}' \\ -\varepsilon_s\Omega_{sy} & \varepsilon_s\Omega_{sx} & -i\omega \end{bmatrix} \quad \text{and} \quad \mathbf{N} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \dfrac{\mathrm{v}_{0sx}}{i\omega} & \dfrac{\mathrm{v}_{0sy}}{i\omega} & 0 \end{bmatrix}.$$

Recalling the definition of the equilibrium velocity of species s,

$$\mathrm{v}_{0s} = \frac{\mathbf{B}_0\,\phi'}{A_s}$$

where $A_s = \begin{cases} \mu_0 n_0 e\,(1+m_e/m_i) & \text{for electrons} \\ -m_i/m_e\,\mu_0 n_0 e\,(1+m_e/m_i) & \text{ions} \end{cases}$, it becomes clear that the elements of $\mathbf{M}$ which include derivatives of the equilibrium velocity will now contain second derivatives of the field angle. Thus, although symbolically the same as Chapter 6, certain elements of **M** may in fact be quite different, depending on the rate of change of the orientation of the ambient magnetic field.

In order to eliminate the perturbed density ($n_s = n_0 / i\omega v_{sz}'$), we need to know v_{sz}' and so require an expression for the third rows of $\mathbf{M}^{-1}$ and $\mathbf{M}^{-1}\cdot\mathbf{N}$, since

$$v_{sz}' = \frac{q_s}{m_s}\left[(\mathbf{M}^{-1})'\mathbf{E} + (\mathbf{M}^{-1} + (\mathbf{M}^{-1}\cdot\mathbf{N})')\mathbf{E}' + \mathbf{M}^{-1}\cdot\mathbf{N}\mathbf{E}''\right]_z .$$

We must therefore calculate the determinant of **M**. Now, although $\phi'' \neq 0$, the second derivatives cancel as before, leaving:

$$\det \mathbf{M} = i\omega(\omega^2 - \Omega_s^2 + v_{0s}\varepsilon_s\Omega_s\phi').$$

There will, however, be a change in the elements of the matrices $(\mathbf{M}^{-1})'$ and $(\mathbf{M}^{-1}\cdot\mathbf{N})'$ through the *derivative* of the determinant of **M** . These changes will result in alterations to the elements of the tensors relating the current density and electric field (cf. equation (6.14)), hence modifying the components of the wave equation, (6.15). Instead of repeating the appropriate tensors in full, we will write the wave equation as:

$$\nabla\times(\nabla\times\mathbf{E}) = \frac{\omega^2}{c^2}(\varepsilon+\varepsilon_1)\cdot\mathbf{E} + i\omega\mu_0((\tau+\tau_1)\cdot\mathbf{E}' + \rho\cdot\mathbf{E}''), \qquad (7.1)$$

where ε, τ and ρ are defined in §6 of Chapter 6. ε_1 and τ_1, which contain the additional terms, are as follows:

$$\varepsilon_1 = \begin{bmatrix} (2T_4 - T_5)\mu\nu & -(2T_4 - T_5)\mu^2 & i\frac{\mu}{\omega}(T_2\frac{\phi''}{\phi'} - 2T_6) \\ (2T_4 - T_5)\nu^2 & -(2T_4 - T_5)\mu\nu & i\frac{\nu}{\omega}(T_2\frac{\phi''}{\phi'} - 2T_6) \\ 0 & 0 & 0 \end{bmatrix},$$

where the ϕ'' terms have been grouped together in T_4, T_5 and T_6, defined by:

$$T_4 = \sum_s \frac{\omega_{ps}^2\Omega_s^2 v_{0s}^2\phi''}{\omega^2(\omega^2 - \Omega_s^2 + \varepsilon_s\Omega_s v_{0s}\phi')^2}, \quad T_5 = \sum_s \frac{\omega_{ps}^2\varepsilon_s\Omega_s v_{0s}\phi''}{\omega^2\phi'(\omega^2 - \Omega_s^2 + \varepsilon_s\Omega_s v_{0s}\phi')},$$

$$T_6 = \sum_s \frac{\omega_{ps}^2\varepsilon_s\Omega_s v_{0s}^2\phi''}{(\omega^2 - \Omega_s^2 + \varepsilon_s\Omega_s v_{0s}\phi')^2}.$$

It must be reiterated that, although these new quantities appear to be multiples of each other, they are not. The summation over species forces us to separate the individual combinations of equilibrium velocities, plasma and cyclotron frequencies into different groupings. Two further definitions are required for τ_1:

$$T_7 = \sum_s \frac{\omega_{ps}^2 v_{0s}^2 \phi''}{\omega\phi'(\omega^2 - \Omega_s^2 + \varepsilon_s \Omega_s v_{0s}\phi')} \quad , \quad T_8 = \sum_s \frac{\omega_{ps}^2 \varepsilon_s \Omega_s v_{0s}^3 \phi''}{\omega(\omega^2 - \Omega_s^2 + \varepsilon_s \Omega_s v_{0s}\phi')^2},$$

allowing τ_1 to be simplified in terms of these new tensor elements to

$$\tau_1 = \frac{\varepsilon_0}{i} \begin{bmatrix} 2(T_7 - T_8)\mu^2 & 2(T_7 - T_8)\mu\nu & 0 \\ 2(T_7 - T_8)\mu\nu & 2(T_7 - T_8)\nu^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We may rearrange the third component of the vector wave equation (7.1) to eliminate E_z in favour of the other two electric field components, using:

$$E_z = -\frac{1}{\varepsilon_{33}}\left(\frac{1}{\omega}((\tau+\tau_1)_{31}E_x' + (\tau+\tau_1)_{32}E_y') + (\varepsilon_{31}E_x + \varepsilon_{32}E_y)\right), \tag{7.2}$$

where we have removed the factor ε_0/i from the elements of $\tau+\tau_1$. Such a straightforward elimination is only possible because the 3-3 elements of $\tau+\tau_1$ and ρ are zero, so that equation (7.2) does not contain any derivatives of the z-component of the electric field. Equation (7.2) must now be differentiated and the right hand side of this expression substituted for E_z' in both remaining components of the wave equation. Two more terms enter into the analysis at this stage, arising from the derivatives of the elements of the modified dielectric tensor, ε :

$$\tilde{S}' = 2\sum_s \frac{\omega_{ps}^2 \varepsilon_s \Omega_s v_{0s} \phi''}{(\omega^2 - \Omega_s^2 + \varepsilon_s \Omega_s v_{0s}\phi')^2} \stackrel{\partial}{=} 2T_9$$

and

$$\tilde{D}' = -2\sum_s \frac{\omega_{ps}^2 \Omega_s^2 v_{0s} \phi''}{\omega(\omega^2 - \Omega_s^2 + \varepsilon_s \Omega_s v_{0s}\phi')^2} \stackrel{\partial}{=} -2T_{10}.$$

Fortunately, when we construct the coupled equations in E_x and E_y, we find that the T terms always occur together in one of two patterns. We may thus define two final quantities to encompass all the terms of order ϕ'', in the same way that we defined Λ, χ and Ξ to include all the ϕ' terms:

$$U_1 = \frac{\omega}{c^2}\left[\frac{T_2^2}{\omega\tilde{S}}\left(\frac{\phi''}{\phi'} - \frac{2T_9}{\tilde{S}}\right) - \frac{4T_2T_6}{\omega\tilde{S}} + 2(T_7 - T_8)\right],$$

$$U_2 = \frac{\omega^2}{c^2}\left[2T_4 - T_5 + \frac{\tilde{D}}{\omega\tilde{S}}\left(2T_6 - T_2\frac{\phi''}{\phi'}\right) + \frac{2T_2}{\omega\tilde{S}}\left(\frac{\tilde{D}T_9}{\tilde{S}} + T_{10}\right)\right].$$

The two coupled, second order ode's in E_x and E_y which are equivalent to equations (6.19) and (6.20) are then of the form:

$$(1-\Lambda\chi\mu^2)E_x''+(2\phi'\Lambda\chi\mu\nu+U_1\mu^2)E_x'+(P\mu^2+\frac{\tilde{R}\tilde{L}}{\tilde{S}}\nu^2-\phi'\chi\mu^2+\Xi\nu^2+U_2\mu\nu)E_x$$

$$=\Lambda\chi\mu\nu E_y''-(\chi-\phi'\Lambda\chi(\mu^2-\nu^2)+U_1\mu\nu)E_y'+(-\mu\nu(P-\frac{\tilde{R}\tilde{L}}{\tilde{S}}-\phi'\chi-\Xi)+U_2\mu^2)E_y \quad (7.3)$$

and

$$(1-\Lambda\chi\nu^2)E_y''+(-2\phi'\Lambda\chi\mu\nu+U_1\nu^2)E_y'+(P\nu^2+\frac{\tilde{R}\tilde{L}}{\tilde{S}}\mu^2-\phi'\chi\nu^2+\Xi\mu^2-U_2\mu\nu)E_y$$

$$=\Lambda\chi\mu\nu E_x''+(\chi+\phi'\Lambda\chi(\mu^2-\nu^2)-U_1\mu\nu)E_x'+(-\mu\nu(P-\frac{\tilde{R}\tilde{L}}{\tilde{S}}-\phi'\chi-\Xi)+U_2\nu^2)E_x, \quad (7.4)$$

where we have absorbed the factor ω^2/c^2 into P and $\tilde{R}\tilde{L}/\tilde{S}$. The terms U_1 and U_2 occur in (7.3) and (7.4) in such a way that the symmetry of these equations has been maintained from Chapter 6. It is therefore possible to define $E_+ = E_x + iE_y$ and $E_- = E_x - iE_y$ in the usual manner and to construct a pair of equations in these variables corresponding to (6.22) and (6.23). Adding equation (7.3) to the product of i and equation (7.4) gives:

$$(1-\frac{\Lambda\chi}{2})E_+''-(i\chi-\frac{U_1}{2})E_+'+(\kappa-\frac{\phi'\chi}{2}+\frac{\Xi}{2}+i\frac{U_2}{2})E_+$$

$$=(\frac{\Lambda\chi}{2}E_-''+(i\frac{\phi'\Lambda\chi}{2}-\frac{U_1}{2})E_-'-(\delta-\frac{\phi'\chi}{2}-\frac{\Xi}{2}-i\frac{U_2}{2})E_-)e^{2i\phi}. \quad (7.5)$$

Subtracting the product of i times equation (7.4) from equation (7.3) yields:

$$(1-\frac{\Lambda\chi}{2})E_-''+(i\chi+\frac{U_1}{2})E_-'+(\kappa-\frac{\phi'\chi}{2}+\frac{\Xi}{2}-i\frac{U_2}{2})E_-$$

$$=(\frac{\Lambda\chi}{2}E_+''-(i\frac{\phi'\Lambda\chi}{2}+\frac{U_1}{2})E_+'-(\delta-\frac{\phi'\chi}{2}-\frac{\Xi}{2}+i\frac{U_2}{2})E_+)e^{-2i\phi}. \quad (7.6)$$

Clearly, the general form of these equations has not been altered, the symmetry has been preserved and the exponential factors have been separated out as before. The situation does not appear to have changed from the comparable stage of Chapter 6. It is, however, at the subsequent stage of the analysis that our method of solution fails. In order to cancel the exponent terms $e^{2i\phi}$ and $e^{-2i\phi}$, we define $E_+ = me^{i\phi}$ and $E_- = ne^{-i\phi}$, but differentiating these substitutions twice will now lead to additional terms, since $\phi'' \neq 0$. The coefficients of the resulting equations, unlike (6.26) and (6.27), will therefore be position dependent although the external phase variation ($e^{\pm 2i\phi}$) has been removed. This remaining spatial dependence makes this problem fundamentally different from the one which we were able to tackle successfully before. Even if ϕ'' were constant, ϕ' would vary and a simple solution using a Fourier transformation would be impossible. The transformation of variables which proved crucially important in Chapter 6 has not made the pair of equations (7.5) and (7.6) any easier to solve than the set (7.3) and (7.4). We conclude that an alternative solution method must be sought for this problem.

4. Non-perpendicular rotation - $k_x \neq 0$

The final extension to the model of Chapter 6 which we will consider is the possibility of waves propagating at other than 90° to the plane of the magnetic field. Suppose we allow a nonzero component of the wavenumber in the x-direction so that the wave is no longer purely perpendicular. The continuity equation will be altered because the gradient of the perturbed fields will now contain the factor k_x, being of the form:

$$\frac{\partial n_s}{\partial t} + \mathbf{v}_{0s}\cdot\nabla n_s + n_0\nabla\cdot\mathbf{v}_s = 0,$$

where n_s, $\mathbf{v}_s$ represent the perturbed density and velocity respectively while n_0 is the equilibrium density common to both species. Fourier transforming in time and the spatially-uniform x-direction yields:

$$-i\omega n_s + \mathrm{v}_{0sx} ik_x n_s + n_0(ik_x \mathrm{v}_{sx} + \mathrm{v}_{sz}') = 0,$$

which may be rearranged to express the perturbed density in terms of the velocity:

$$n_s = \frac{n_0(ik_x \mathrm{v}_{sx} + \mathrm{v}_{sz}')}{i(\omega - k_x \mathrm{v}_{0sx})}. \tag{7.7}$$

Note that this expression contains two new terms which depend on the horizontal component of the wavenumber. The momentum equation:

$$m(\frac{\partial \mathbf{v}_s}{\partial t} + \mathbf{v}_{0s}\cdot\nabla\mathbf{v}_s + \mathbf{v}_s\cdot\nabla\mathbf{v}_{0s}) = q_s(\mathbf{E} + \mathbf{v}_{0s}\times\mathbf{B}_1 + \mathbf{v}_s\times\mathbf{B}_0),$$

will be altered similarly and may be expressed more clearly in tensor notation as:

$$\mathbf{M}\mathbf{v}_s = \frac{q_s}{m_s}(\mathbf{I}+\mathbf{P})\,\mathbf{E} + \frac{q_s}{m_s}\mathbf{N}\mathbf{E}',$$

where

$$\mathbf{M} = \begin{bmatrix} -i\omega + ik_x \mathrm{v}_{0sx} & 0 & \varepsilon_s\Omega_{sy} + \mathrm{v}_{0sx}' \\ 0 & -i\omega + ik_x \mathrm{v}_{0sx} & -\varepsilon_s\Omega_{sy} + \mathrm{v}_{0sy}' \\ -\varepsilon_s\Omega_{sy} & \varepsilon_s\Omega_{sx} & -i\omega + ik_x \mathrm{v}_{0sx} \end{bmatrix}, \quad \mathbf{P} = \frac{k_x}{\omega}\begin{bmatrix} 0 & \mathrm{v}_{0sy} & 0 \\ 0 & -\mathrm{v}_{0sx} & 0 \\ 0 & 0 & -\mathrm{v}_{0sx} \end{bmatrix}$$

and **N** is unchanged from §3.

From equation (7.7) we see that, in order to eliminate the number density, we will now require, not only the derivative of the z component of the first-order velocity vector, but also its x component. The determinant of the matrix **M** is:

$$\det \mathbf{M} = i\,(\omega - k_x \mathrm{v}_{0sx})((\omega - k_x \mathrm{v}_{0sx})^2 - \Omega_s^2 + \varepsilon_s \Omega_s \mathrm{v}_{0s} \phi')$$

and so we see that the frequency has undergone an effective Doppler shift: $\omega \to \omega - k_x \mathrm{v}_{0sx}$. When the current density equation,

$$\mathbf{J} = \sum_s q_s (n_0 \mathbf{v}_s + n_s \mathbf{v}_{0s}),$$

is evaluated in the usual way, the denominators of all the tensor elements will be changed accordingly. The 3-3 elements of the tensors τ and ρ (which multiply E_z' and E_z'' respectively) will be zero, as in the cases considered previously. The third row of ρ is composed solely of the third row of the tensor $\mathbf{M}^{-1}{\cdot}\mathbf{N}$ multiplied by the z-component of the equilibrium velocity, which is zero by definition. Similarly, the last row of τ contains a product with v_{0sz} which vanishes and the remaining contribution to the 3-3 element of τ comes from $(\mathbf{M}^{-1}{\cdot}\mathbf{N})_{3,3}$ which vanishes because the final column of **N** is zero. Clearly, the elimination of E_z will therefore be trivial, and the system could be expressed as a pair of coupled, second order ode's in E_x and E_y as before.

We will not proceed beyond this point in the analysis since it is apparent that, ultimately, it will not be possible to reduce the differential equations to ones with constant coefficients. Even if $\phi'' = 0$, the coefficients of the differential equations will not be constant since they will include, at least, the factors $k_x \mathrm{v}_{0sx} = \mu k_x \mathrm{v}_{0s}$ and its derivative, $-\nu k_x \mathrm{v}_{0s}$, which are both position-dependent. A further problem remains with the solution technique used before. The asymmetry of the system caused by the introduction of the horizontal wavenumber component is likely to abrogate the usefulness of the transformation of the electric field components from E_x and E_y to E_+ and E_-. All in all, another method of solution will be required before the case of non-perpendicular propagation may be solved completely.

5. Conclusions

One popular solution method for coupled differential equations is that used by Cairns and Lashmore-Davies (1983). The first of the pair of differential equations is solved for one dependent variable, assuming that the coupling is negligible so that the "right hand side" of the equation may be set equal to zero. This solution is then substituted into the right hand side of the second equation where it acts as a driving term when this equation is solved. Finally, the resulting solution for the second dependent variable is resubstituted into the original equation. The accuracy of this method may be improved by successive iterations of this procedure which is therefore easily performed by a computer. This method, of course relies on the equations being soluble in the first instance. It is also entirely unsuitable for the case in which the coefficients are periodic since it would be impossible to make any assumption of negligible coupling while retaining identical terms on the other side of the equation. Equally, periodic coefficients can only be small in short sections of the plasma and will grow again until they cannot be ignored so that solutions valid for small coefficients are invalid over most of the plasma.

There are clearly formidable problems to be faced in solving wave propagation problems for inhomogeneous media. The acoustic-gravity waves of Chapter 4 were described by a second order equation which, although apparently simple, did not possess a standard solution. Many investigations of mode conversion and related topics have faltered at this hurdle. Often, the method has been to approximate the differential equation to one with an analytic solution, such as the Weber equation, but it is optimistic to assume that all problems of wave propagation can be described by this one equation. The limited number of differential equations with analytic solutions is a major difficulty in studies of wave motion. The alternative approach, via position-dependent dispersion relations was demonstrated graphically in Chapter 4 to be of limited validity.

In this chapter, we have considered two minor extensions to the model of Chapter 6 (which was solved completely and analytically) and have discovered that neither can be solved by similar methods. Had we attempted to model the exact magnetic field configuration of a tokamak or a region of the solar atmosphere, say, we would have been forced to incorporate even greater complexity and consequently would have had even less chance of deriving an exact solution. It was demonstrated in §3.1 of Chapter 6 that, under our equilibrium assumptions, the ambient magnetic field could not have a component in the direction of variation if this variation were to be expressible through a single independent variable. Thus, partial differential equation descriptions would become necessary for any more complicated magnetic field topology. It seems inevitable that progress can only be made in the areas described via some kind of approximation technique, but the results of any such approximation should be weighed carefully against the evidence available from analytic solutions of simpler cases.

This thesis has attempted to convey two principal ideas. First, local dispersion relation techniques cannot be relied upon to describe adequately the conditions in an inhomogeneous medium - cf. Chapters 2 and 4. Second, physical intuition is not sufficient to construct a physical model of an inhomogeneous medium - see, in particular, the early sections of Chapter 6. An accurate description of the propagation of waves in inhomogeneous media is only possible through a compromise which combines a rigorous mathematical description with a thorough physical understanding. Neither can be wholly successful without the other and, because of the limitations of both, advances in this subject can only be made if they are used wisely.

Appendix - Extension of eigenvalue analysis to fourth order ode's

In order to prove that the theory of Clemmow and Heading (1954) also applies to the case of fourth order ode's, let us consider a general fourth order ode with spatially varying coefficients:

$$a(z)y^{iv}+b(z)y'''+c(z)y''+d(z)y'+e(z)y=0. \tag{A.1}$$

In exactly the same way as equation (2.1) was written in vector notation, we may write (A.1) as:

$$\mathbf{y}'=\mathbf{My}=\mathbf{M}\begin{bmatrix} y \\ y' \\ y'' \\ y''' \end{bmatrix}, \tag{A.2}$$

where we now have

$$\mathbf{M}=\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -e/a & -d/a & -c/a & -b/a \end{bmatrix}.$$

The eigenvalues of this matrix, which we again consider to be the eigenvalues of equation (A.1), λ_1, λ_2, λ_3, λ_4 are the solutions of the polynomial

$$\lambda^4+\frac{b}{a}\lambda^3+\frac{c}{a}\lambda^2+\frac{d}{a}\lambda+\frac{e}{a}=0.$$

A suitable diagonalising matrix, **A**, can be constructed from the column eigenvectors corresponding to each eigenvalue where the ith column of **A** is given by :

$$\begin{bmatrix} 1 \\ \lambda_i \\ \lambda_i^2 \\ \lambda_i^3 \end{bmatrix}.$$

Thus, when the transformation $\mathbf{y}=\mathbf{Au}$, is performed, the system of equations represented by (A.2) becomes:

$$\mathbf{u}' = \mathbf{A}^{-1}\mathbf{MAu} - \mathbf{A}^{-1}\mathbf{A}'\mathbf{u},$$

where the first matrix on the right hand side of this equation is, by definition, a diagonal matrix composed of the eigenvalues of **M**. The second matrix which represents the coupling has more complicated elements than its counterpart in two dimensions but maintains the symmetry about the main diagonal observed in the two-dimensional case. A diagonal element is defined by:

$$(\mathbf{A}^{-1}\mathbf{A}')_{ii} = \lambda_i'(\frac{1}{\lambda_i - \lambda_j} + \frac{1}{\lambda_i - \lambda_k} + \frac{1}{\lambda_i - \lambda_l}),$$

where $i, j, k, l \in (1,..,4)$ and an off diagonal *cross-coupling* term will have the form:

$$(\mathbf{A}^{-1}\mathbf{A}')_{ij} = \lambda_j' \frac{(\lambda_j - \lambda_k)(\lambda_j - \lambda_l)}{(\lambda_i - \lambda_j)(\lambda_i - \lambda_k)(\lambda_i - \lambda_l)}.$$

We notice immediately that by setting $\lambda_{3,4}$ equal to zero we retrieve the coupling matrix for two dimensions, derived in Chapter 2. The off-diagonal terms which determine the coupling between different modes share the same dependences which were observed for the 2×2 case. The larger the variation of an eigenvalue, the more strongly does it interact. Another attribute of the coupling which seems physically reasonable is that the coupling of, for instance, λ_2 to λ_1 increases as these two eigenvalues become closer together and also as λ_2 gets further away from the remaining two eigenvalues.

The form of the coupling matrix in the 3×3 case may be deduced easily from the definitions of the A_{ij}. Similarly, it would seem that this system of identification of modes should apply to all higher orders of equation with the interaction between modes being regulated in a similar manner and we postulate that the coefficients would have the general form:

$$A_{ii} = \lambda_i' \sum_{k \neq i} \frac{1}{\lambda_i - \lambda_k} \quad \text{and} \quad A_{ij} = \lambda_j' \frac{\prod_{k \neq i,j} (\lambda_j - \lambda_k)}{\prod_{l \neq i} (\lambda_i - \lambda_l)}.$$

References

ADAM J.A. (1977) *Solar Phys.*, **52**, 293.

ALFVEN H. (1942) *Arkiv. Mat. Astron. Fysik.*, **29B** No. 2.

ALLIS W.P (1959) *Mass. Inst. Technol. Research Lab. Electronics Quart. Prog. Rept.*, **54**, 5.

APPLETON E.V. (1932) *J. Inst. Elec. Engrs.*, **71**, 642.

ANDO H. and OSAKI Y. (1977) *Publications of the Astronomical Society of Japan*, **29**, 221.

ASTROM E.O. (1950) *Arkiv. Fysik*, **2**, 443.

BATEMAN MANUSCRIPT PROJECT STAFF (1953) ed. Erdelyi *Higher Transcendental Functions*, McGraw-Hill, New York.

BEER T. (1974) *Atmospheric waves*, Adam Hilger, London.

BEL N. and MEIN P. (1971) *Astron. Ap.*, **11**, 234.

BENNETT W.M. (1934) *Phys. Rev.*, **45**, 890.

BLOCH F (1928) *Z. Physik*, **52**, 555.

BOYD T.J.M. and SANDERSON J.J. (1969) *Plasma Dynamics*, Nelson, London.

BRILLOUIN L. (1926) *C. R. Acad. Sci. Paris*, **183**, 24.

BRILLOUIN L. (1953) *Wave Propagation in Periodic Structures*, Dover, New York.

BRUNT D. (1927) *Quart. J. Roy. Meteor. Soc.*, **53**, 30.

BUDDEN K.G. (1961) *Radio Waves in the Ionosphere*, Cambridge Univ. Press.

CAIRNS R.A. and LASHMORE-DAVIES C.N. (1982) *Phys. Fluids*, **25**, 1605.

CAIRNS R.A. and LASHMORE-DAVIES C.N. (1983) *Phys. Fluids*, **26**, 1268.

CASSELS J.M. (1982) *Basic Quantum Mechanics*, Macmillan Press, London.

CHANDRASEKHAR S. and WOLTJER L. (1958) *Proc. Nat. Acad. Sci.*, **44**, 285.

CHAPMAN S. and COWLING T.G. (1960), *The Mathematical Theory of Nonuniform Gases*, 2nd ed., Cambridge University Press, London.

CHEN F.F. (1984) *Introduction to Plasma Physics and Controlled Fusion,* Plenum Press, New York.

CHIMONAS G. and HINES C.O. (1986) *J. Geophys. Res.,* **91**, 1219.

CHOUDHURY S.R. (1988) *Phys. Fluids,* **31**, 2206.

CHRISTENSEN-DALSGAARD J. (1980) *Mon. Not. R. astr. Soc.,* **190**, 765.

CIRA (1965) *Cospar international reference atmosphere*, North Holland Publ. Co., Amsterdam.

CIRA (1972) *Cospar international reference atmosphere*, Akademie-Verlag, Berlin.

CLEMMOW P.C. and HEADING J. (1954) *Proc. Camb. Phil. Soc.,* **50**,391.

DANIELS G.M. (1967) *J. Geophys. Res.,* **72**, 2419.

DIVER D.A. (1986) *Ph.D. Thesis* , University of Glasgow.

DIVER D.A. (1986) *8th Europhysics Conference on Computational Physics: Computing in Plasma Physics, Eibsee,* **139**.

DIVER D.A. and LAING E.W. (1989) submitted to *J. Plasma Phys.*.

DURRANT C.J. (1988) *The Atmosphere of the Sun*, Adam Hilger, London.

ECKART C. (1960) *Hydrodynamics of Oceans and Atmospheres*, Pergamon, London.

EINAUDI F. and HINES C.O. (1969) *Can. J. Phys.,* **48**, 1458.

FIDONE I. and GRANATA G. (1971) *Nucl. Fusion,* **11**, 133.

FLOQUET G. (1883) *Ann. de l'Ecole norm. sup.,* (2), XII.

FRANCIS S.H. (1975) *J. Atmos. Terr. Phys.,* **37**, 1101.

FRIEDLAND L. (1986) *Phys. Fluids,* **29**, 1105.

FRISCH U. (1964) *Ann. Astrophys.,* **27**, 224.

FUCHS V., KO K. and BERS A. (1985) *Phys. Fluids,* **28**, 177.

GOSSARD E.E. and HOOKE W.H. (1975) *Waves in the Atmosphere,* Elsevier, Amsterdam.

GROSSMAN W. and WEITZNER H. (1984) *Phys. Fluids,* **27**, 1699.

HARTREE D.R. (1931) *Proc. Cambridge Phil. Soc.,* **27**, 143.

HEADING J. (1962) *An Introduction to Phase Integral Methods*, Methuen, London.

HEWISH A. *et al.* (1968) *Nature,* **217**, 709.

HINES C.O. (1960) *Can. J. Phys.,* **38**, 1441.

HOYAUX M.F. (1970) *Solid State Plasmas,* Pion, London.

JEFFREYS H. (1923) *Proc. Lond. Math. Soc.,* **23**, 428.

JETTE D.A. (1970) *J. Math. An. and Appl.,* **29**, 109.

JONES W.L. (1970) *J. Atmos. Terr. Phys.,* **32**, 1555.

JORDAN S. (ed.) (1981) *The Sun as a Star,* **NASA SP-450** , Washington/ Paris:NASA/CNRS.

KRAMERS H.A. (1926) *Zeit. f. Phys.,* **39**, 828.

KRISTIAN J. et al. (1989) *Nature,* **338**, 234.

LAMB H. (1945) *Hydrodynamics,* 6th. ed., Dover, New York.

LANDAU L.D. and LIFSCHITZ E.M. (1959) *Fluid Mechanics* , Pergamon Press, Oxford.

LASHMORE-DAVIES C.N. and STENFLO L. (1981) *Phys. Fluids,* **24**, 984.

LAWSON J.D. (1957) *Proc. Phys. Soc. B,* **70**, 6.

LEIGHTON R.B. (1960), Communication to the 4th I.A.U.-I.U.T.A.M., *Symposium on Cosmical Gas Dynamics.*

LEIGHTON R.B., NOYES R.W. and SIMON G.W. (1962) *Ap. J.,* **135**, 474.

MARTYN D.F. (1950) *Proc. Roy. Soc.,* **A201**, 216.

MASSEY H.S.W. (1964) *Space Physics,* Cambridge University Press, London.

MATHIEU E.(1868) *Journal de Math.* , (2), XIII.

McLACHLAN N.W. (1947) *Theory and Application of Mathieu functions* , Clarendon Press, Oxford.

McLELLAN A. and WINTERBERG F. (1968) *Solar Phys.,* **4**, 401.

MELROSE D.B. (1974a) *Aust. J. Phys.,* **27**, 31.

MELROSE D.B. (1974b) *Aust. J. Phys.,* **27**, 43.

MELROSE D.B. (1977a) *Aust. J. Phys.,* **30**, 495.

MELROSE D.B. (1977b) *Aust. J. Phys.,* **30**, 661.

MELROSE D.B. (1980) *Aust. J. Phys.,* **33**, 121.

MELROSE D.B. and SIMPSON M.A. (1977) *Aust. J. Phys.,* **30**, 647.

MIKUMO T. (1968) *J. Geophys. Res.,* **73**, 2009.

MOORE D.W. and SPIEGEL E.A. (1964) *Ap. J.,* **139**, 48.

MORSE P.M. and FESHBACH H. (1953) *Methods of Theoretical Physics*, McGraw-Hill, New York.

MURPHY G.M. (1960) *Ordinary Differential Equations and their Solutions*, Van Nostrand, New York.

NAKAGAWA Y. and RAADU M.A. (1972) *Solar Phys.,* **25**, 127.

OSTERBROCK D.E. (1961) *Ap. J.,* **134**, 347.

OSTRIKER J.P. and GUNN J.E. (1969) *Astrophys. J.,* **157**, 1395.

PEKERIS C.L. (1939) *Proc. R. Soc. Lond.,* **A171**, 434.

PEKERIS C.L. (1948a) *Phys. Rev. 2nd Ser.,* **73**, 145.

PENNING F.M. (1926) *Physica,* **6**, 241.

PITTEWAY M.L.V. and HINES C.O. (1965) *Can. J. Phys.,* **43**, 2222.

REDUCE 3.3 User's Manual (1987) High Level Hardware Ltd., Windmill Road, Oxford.

SCORER R.S. (1950) *Proc. Roy. Soc.,* **A201**, 137.

SHKLOVSKY I.S. (1946) *Astron. Zh. USSR,* **23**, 333.

SITENKO A.G. and STEPANOV K.N. (1957), *Soviet Phys. JETP,* **4**, 512.

SNEDDON I.N. (1980) *Special Functions of Mathematical Physics and Chemistry*, Longman Group Ltd., London.

STEIN R.F. and LEIBACHER J. (1974) *Ann. Rev. Astron. Astrophys.,* **12**, 407.

STIX T. (1962) *The Theory of Plasma Waves*, McGraw-Hill, New York.

STIX T. (1965) *Phys. Rev. Lett.,* **15**, 878.

STIX T. and SWANSON D.G. (1983) *Handbook of Plasma Physics*, ed. A. Galeev and R.N. Sudan, North-Holland, Amsterdam, Vol. 1.

SWANSON D.G. (1985), Review Article: *Phys. Fluids,* **28**, 2645.

TAYLER R.J. (1982) *Plasma Physics and Nuclear Fusion Research* , ed. R.D. Gill, Academic Press, London.

TAYLOR J.B. (1974) *Phys. Rev. Lett.,* **33**, 1139.

TER HAAR D., *Elements of Statistical Mechanics* , Holt, Rinehart and Winston Inc., New York, 1954.

TONKS L. and LANGMUIR I. (1929) *Phys. Rev.,* **33**, 195.

TUAN T.F. and TADIC D. (1982) *J. Geophys. Res.*, **87**, 1648.

UCHIDA Y. (1965) *Ap. J.*, **142**, 335.

VAISALA V. (1925) *Soc. Scient. Fennica Comm. Phys.-Math II,* **8**, 1.

WENTZEL G. (1926) *Zeit. f. Phys.*, **38**, 518.

WHITTAKER E. and WATSON G.N. (1927) *A Course of Modern Analysis* , 4th ed., Cambridge University Press, New York.

WOLTJER L. (1958) *Proc. Nat. Acad. Sci.,* **44**, 489.

YU L., TUAN T.F. and TAI H. (1980) *J. Geophys. Res.*, **85**, 1297.

ZHUGHZDA Yu. D. and DZHALILOV N.S. (1984) *Astron. Astrophys.,* **132**, 45.

ZHUGHZDA Yu. D. and DZHALILOV N.S. (1984) *Astron. Astrophys.,* **132**, 52.

ZHUGHZDA Yu. D. and DZHALILOV N.S. (1984) *Astron. Astrophys.,* **133**, 333.