THE
MAGNETO-IONIC THEORY
AND ITS APPLICATIONS
TO THE IONOSPHERE
A MONOGRAPH

BY
J. A. RATCLIFFE
C.B.E., M.A., F.R.S.
Director of the Radio Research Station, D.S.I.R.

CAMBRIDGE
AT THE UNIVERSITY PRESS
1962
CONTENTS

Preface

1. Introduction

1.1 Purpose of the book
1.2 Plan of the book
1.3 Bibliography
1.4 History
1.5 Nomenclature

PART I. THE DERIVATION OF THE EQUATIONS

2. The macroscopic approach. Derivation of Appleton's equations by means of an equivalent continuous medium

2.1 The equivalent continuous magneto-ionic medium
8
2.2 A wave in a continuous medium of dielectric constant \( \varepsilon \)
11
2.3 Waves in a medium which is electrically non-isotropic. Characteristic waves
12
2.4 The constitutive relations for a magneto-ionic medium
14
2.5 The characteristic waves in the magneto-ionic medium
16
2.6 The effects of collisions
18

3. The microscopic approach. The physical reason for the characteristic waves and their dispersion

3.1 The microscopic explanation of dispersion
20
3.2 The wave scattered from an element of volume
22
3.3 Scattering from a thin slab
22
3.4 Effect of a thin slab of polarisable material on waves passing normally through it
25
3.5 Propagation through a continuous medium when the polarisability may be large. Fundamental concepts
26
3.6 Calculation of the self-consistent wave
27
3.7 Discussion of the self-consistent wave
29
3.8 The characteristic waves of the magneto-ionic theory
29
CONTENTS

4. Absorption on the ‘macroscopic’ theory
   4.1 Introduction
   4.2 The frequency of collision of an electron with heavy particles
   4.3 The average effect of collisions on the motion of an electron
   4.4 The absorption coefficient in the magneto-ionic theory

5. The microscopic picture of absorption
   5.1 Introduction
   5.2 Absorption in terms of energy-transfer
   5.3 Extension to the case where a magnetic field is present (longitudinal case)
   5.4 Absorption in terms of re-radiation from individual electrons

PART II. INTERPRETATION OF THE EQUATIONS

6. Graphical representation when there are no collisions
   6.1 No imposed magnetic field
   6.2 The longitudinal case
   6.3 The transverse case
   6.4 The general case (Y less than unity, e.g. Y = 1)
   6.5 The general case (Y greater than unity, e.g. Y = 2)
   6.6 Ordinary and Extraordinary waves

7. The polarisations of the waves
   7.1 Introduction
   7.2 The parameter $\omega_c$ and the polarisation when $X = 1$
   7.3 Ordinary and Extraordinary waves and the choice of sign in the equations
   7.4 Curves of $R(X)$
   7.5 Limiting polarisation

8. The quasi-longitudinal (QL) and quasi-transverse (QT) approximations
   8.1 The approximations
   8.2 Conditions which satisfy the approximations when $X = 1$
   8.3 The two approximations when $X \neq 1$
      (a) The QT approximation
      (b) The QL approximation
   8.4 The QL and QT regions in a model ionosphere
   8.5 The relation between the QL expression and that for the case of no applied field

9. The effect of collisions on the curves for $\mu$ and $\chi$
   9.1 Introduction
   9.2 Curves for $\mu^{\prime}(X)$ and $\chi^{\prime}(X)$ (general)
   9.3 The detailed changes of $\mu$ and $\chi$
      (a) The curves for $\mu^{\prime}(X)$
      (b) The curves for $\chi^{\prime}(X)$
   9.4 Approximations for small values of $Z$, and the conditions for reflection
   9.5 Conditions for reflection when $Z^2 \gg 1$ at $X = 1$
   9.6 Results for the QL approximation
   9.7 The QT approximation near $X = 1$, upper sign

10. Some representative curves for the refractive index ($\mu$) and the absorption index ($\chi$)
   10.1 Introduction
   10.2 The magnitudes of $\mu$ and $\chi$ near the possible reflection levels
   10.3 The method for sketching approximate curves
   10.4 Some particular examples

11. The group velocity
   11.1 No magnetic field
   11.2 An imposed magnetic field
CONTENTS

Part III. Applications to the Terrestrial Ionosphere

12. A model ionosphere with no magnetic field
   12.1 Two model ionospheres
   12.2 The penetration frequency
   12.3 The height of reflection
   12.4 The equivalent height of reflection
   12.5 The absorption of the wave
      (a) Non-deviative absorption
      (b) Deviative absorption
   12.6 A model ionosphere with a magnetic field
      13.1 Form of the $h'(f)$ curves over a wide range of frequency
      13.2 Magneto-ionic splitting
      13.3 The effect of collisions on magneto-ionic splitting
      13.4 Shapes of the $h'(f)$ curves for high frequencies
      13.5 Triple splitting and the Z-trace
      13.6 Non-deviative absorption, $QL$ approximation
      13.7 Deviative absorption
         (a) Ordinary wave, $QT$ approximation
         (b) Extraordinary wave, $QL$ approximation
      13.8 Gyro-frequency splitting
      13.9 The 'fourth reflection condition'
      13.10 Low and very low frequencies and the lowest ionosphere
      13.11 $\mu(h)$ and $\chi(h)$ curves for different frequencies in the model ionosphere of fig. 13.13

14. The nature of the electrons' motions in the ionosphere
   14.1 Magnitudes in the absence of a magnetic field
   14.2 Heating of the electrons and ionospheric cross-modulation
      (imposed field neglected)
   14.3 Magnitudes in the presence of a magnetic field
   14.4 Energy relations in the presence of a magnetic field

Part IV. Some Miscellaneous Considerations

15. The Lorentz Term
   15.1 Introduction
   15.2 History
   15.3 Consequences of including the Lorentz Term ($l = \frac{1}{2}$)
   15.4 Experimental test of the Lorentz theory
      (a) Oblique reflection from a layer
      (b) Phenomena near the gyro-frequency
      (c) Very low frequencies
      (d) Results from rocket experiments
16. Waves in a mixture of ions and electrons
   16.1 Introduction
   16.2 The refractive index of a mixture
   16.3 The results for some special cases
   16.4 The effects of collisions in electron-ion mixtures
17. Inhomogeneous media
   17.1 Introduction
   17.2 Conditions for reflection at vertical incidence (no imposed field)
   17.3 Coupling between characteristic waves
   17.4 The coupling echo
   17.5 Emergent polarisation
      (a) Passage through a region of strong coupling
      (b) Limiting polarisation
18. A magneto-ionic medium considered as an anisotropic dielectric, and a comparison with crystal optics
   18.1 Introduction
   18.2 Calculation of refractive index using tensor notation
   18.3 The dielectric and susceptibility tensors
   18.4 The principal directions
PREFACE

In the writing of this book I have had considerable help from numerous friends who have read and commented upon it at various stages. I am particularly indebted to Dr P. C. Clemmow, Dr K. G. Budden and Mr J. P. Dougherty for help of this kind.

For some of the numerical computations of Chapter 10 I am again indebted to Dr K. G. Budden, and to Professor A. H. Waynick and Dr F. Lied for arranging permission for me to quote results from unpublished reports.

Some of the results of Chapter 9 were first obtained by Mrs J. Haselgrove, to whom I am much indebted for permission to include them here.

J.A.R.

Cambridge
August 1958

NOTE ON THE SECOND IMPRESSION

In the process of reprinting the opportunity has been taken to correct some mistakes concerning the signs to be attributed to the two characteristic waves. The nature of the corrections is explained in §6.1 on p. 53.

J.A.R.
CHAPTER I  
INTRODUCTION

1.1. Purpose of the book

This book deals with the theory of electro-magnetic waves passing through a gas of neutral molecules in which is embedded a statistically homogeneous mixture of free electrons and neutralising heavy positive ions, in the presence of an imposed uniform magnetic field. A medium of that kind will be called a magneto-ionic medium and the theory will be called the magneto-ionic theory. It has been mainly applied to problems of radio-wave propagation through the ionosphere.

Although the ionosphere is not, in fact, a homogeneous medium, and a full discussion of the propagation of waves through it requires a full-wave theory of considerable complexity, it is often sufficiently accurate to suppose that the wave behaves, at each level, as though it were propagated in a homogeneous medium. Chapter 17 contains an elementary discussion of the conditions under which this approximation can be made. The rest of the book is concerned only with propagation through a homogeneous magneto-ionic medium and with those problems of radio-wave propagation through the terrestrial ionosphere which can be solved with the help of this theory.

Although much of the book is based on published papers, referred to in the Bibliography, substantial parts of it have not previously been published. They have, in consequence, not been submitted to public criticism as published papers would have been, and the reader is warned to accept them with caution. The sections 3.5, 3.6, 3.7, 3.8, 5.2, 5.3, 5.4, and Chapters 9, 10 and 14 are of this nature.

1.2. Plan of the book

Part I is concerned with the derivation of Appleton’s equations (2.6.6–2.6.10) which express the complex refractive indices and the wave-polarisations for waves travelling in a magneto-ionic medium. The object is to show in detail the physical processes involved in the
wave-propagation and the account is therefore somewhat lengthy and is purposely not in its most concise mathematical form. Two different approaches are used, a macroscopic one (Chapters 2 and 4) in which the properties of the medium are averaged out, and a microscopic one (Chapters 3 and 5) in which attention is concentrated on the motions of the individual electrons. A more concise mathematical formulation is presented in § 18.2.

Appleton's equations are complicated and it is not easy to understand their detailed implications. The object of Part II is to summarise the results in graphical form, demonstrating how the curves showing refractive index and absorption index as a function of electron density can be sketched in outline with the help of some simple rules. A representative series of computed curves is examined in the light of these rules in Chapter 10. It is suggested that, with these curves, and the rules, it is a simple matter to sketch corresponding curves for all possible conditions.

Part III is concerned with the application of the theory to the terrestrial ionosphere. The object has been to include all those phenomena, known to the author, which receive explanation in terms of the magneto-ionic theory of a homogeneous medium.

1.3. Bibliography

The Bibliography contains a list of papers dealing primarily with the magneto-ionic theory or its applications, and a subsidiary list of papers, not essentially theoretical, which are referred to in the text. A summary of the contents of the more important papers is provided. Textual references have been kept to a minimum, but the Bibliography contains detailed references to the individual sections of the text in turn, with notes where necessary.

1.4. History

When, in 1902, Kennelly and Heaviside independently suggested that radio waves might be propagated from England to America, 'round the protuberance of the earth', by reflection from an electrified layer in the upper atmosphere, they both spoke of such a layer as being 'conducting', and they likened its effect to that of the conducting sea water below.

Eccles, in 1912, was the first to consider the matter theoretically, and he derived expressions for the velocity and the absorption of waves passing through a medium containing free charges. His expressions apply only when the velocity of the waves is not much different from that in free space, but with this restriction they were the same as those now in use. The interpretation given to the theory by Eccles was, however, different in two respects from that now given. In the first place he supposed that the effective charges were ions of atomic mass, and in the second place he was imbued with the earlier idea that the medium must act like a conductor. He found it difficult to reconcile these two conditions, since the conducting behaviour required that the collision frequency should be large compared with the angular wave-frequency and it was difficult to see how this could be, particularly with massive charges. Partly because his conducting medium would be strongly absorbing, and partly because he was chiefly concerned with waves of great length, he considered that the waves were returned to the earth by a process of partial reflection at a relatively sharp boundary.

The next theoretical attack on the problem was made by Larmor in 1924. The essential difference between the approaches of Larmor and Eccles is not, as is often said, that Larmor considered electrons and Eccles ions of atomic mass, but that Larmor supposed the collision frequency to be small enough for the medium to behave like a dielectric, while Eccles considered the frequency to be large and the medium like a conductor. A conductor will absorb strongly, and Eccles, realising this, had to postulate a sharply bounded 'Heaviside' layer from which the waves would be partially reflected without being absorbed. In Larmor's dielectric medium, however, the absorption could be much smaller and the wave could be returned to earth by a process of gradual refraction. At the present time Larmor's picture of the ionosphere is thought to be correct for all except the longest waves, for which a mechanism similar to that of Eccles is supposed to be effective (§ 12.3).

In 1925 papers were published almost simultaneously in England and America drawing attention to the importance of the earth's magnetic field in influencing the return of waves from the ionosphere. Both sets of workers had noticed that the propagation of waves of length near 200 m. was unusually poor, and both suggested that this might be because that wavelength corresponded roughly to
the gyro-frequency of free electrons in the earth's magnetic field. Most workers at present would probably consider that this was not the cause of the poor propagation, since, as was pointed out by Appleton (1927), it would affect only the Extraordinary wave, and some other mechanism must be sought to explain it. It is not yet clear what that mechanism is, but it is interesting to notice that, as so often happens in science, a clue wrongly interpreted led to some most interesting results. The English papers emphasising the importance of the magnetic field were published by Appleton and by Appleton & Barnett in February and April 1925, and those from America by Nichols & Schelleng in March (1925 a) and April (1925 b).

Appleton & Barnett referred to Larmor's paper and pointed out the need to take account of the earth's magnetic field. They quoted from Lorentz (1909) the appropriate expressions for the refractive index in the cases of longitudinal and transverse propagation. They were doubtful whether electrons or more massive ions were the effective charged particles, and suggested that a decision could be made by considering the different gyro-frequencies of the two. The effect of collisions was not included in their equations, though there was some discussion of the part they would be expected to play. The name \textit{magneto-ionic theory} was suggested.

Nichols & Schelleng set down the equations for the propagation of a wave in a general direction, and included the effect of collisions, but they did not derive explicit expressions for the refractive index or the absorption coefficient. They showed how to express the equations in terms of a dielectric tensor and referred to Drude's \textit{Optics} (1902) for a solution.

In 1926 Taylor & Hulburt again emphasised the increased absorption on a wave-length near 200 m., suggested that it might be related to the gyro-frequency of free electrons, and quoted Lorentz's results for longitudinal and transverse propagation.

In 1927 Breit considered the propagation of a wave travelling in a general direction relative to the magnetic field, but, like Nichols & Schelleng, he left his results in the form of an unsolved equation to give the magnitude of the refractive index or the absorption coefficient. In the same year Lassen (1927) gave a full analysis for the case of propagation in a general direction and included the possibility of collisions between the electrons and heavy particles. He derived explicit expressions for the refractive index and absorption coefficient, but they do not seem to have been much used, possibly because they were expressed in a somewhat complicated and unfamiliar nomenclature.

The work on which this book is mainly based was published in outline by Appleton, also in 1927. In it he extended the analysis of Lorentz and gave an expression for the complex refractive index, in terms of symbols taken over from the theory of Lorentz. The fact that Appleton's expression has been more widely used than the others appears to indicate that it was more convenient, and possibly that, being based on that of Lorentz, it was more familiar. In 1932 Appleton gave a full account of his treatment and, although he used the symbols of Lorentz in the body of the paper, he introduced, in the Appendix, another set of symbols which have since proved very convenient.

Between 1927, when Appleton published the outline of his theory, and 1932, when he gave it in full, other papers were published on the subject. Goldstein (1928) derived expressions which were essentially the same and applied them to some special problems. Hartree (1931 a) adapted to the present purpose a method of calculation due to Darwin (1924) in which the complex refractive index was deduced by considering the scattered waves re-radiated from the free electrons. This 'microscopic' approach to the problem can be contrasted with the 'macroscopic' approach of the other papers. Hartree included the 'Lorentz Term' (see Chapter 1) which had not been included by the other workers. Burnett (1931) considered the statistics of the electron-collisions in more detail than the other workers.

Since the publication of Appleton's 1932 paper there have been other derivations of the expressions for the complex refractive index, but none has added anything new in principle.

1.5. \textbf{Nomenclature}

The following nomenclature will be used. Harmonically varying quantities will be written $\exp(i\omega t)$. Field Vectors will be printed in heavy type ($\mathbf{E}$, $\mathbf{\mathcal{E}}$, etc.) if their vector nature is of importance, and their scalar components will be written $E_x$, $\mathcal{E}_x$, etc. In general
waves will be supposed to travel along the positive Oz-direction. An attenuated simple harmonic wave-motion, which could be written, for example, like

\[ E = \mathcal{E}_0 e^{-\kappa z} \exp (-ikz) \exp (i\omega t), \]

will be represented from time to time in different ways according to the following scheme, where all the expressions represent the same quantity:

- \( E \)
- \( \mathcal{E}_0 \exp (-ikz) \)
- \( \mathcal{E}_0 \exp (i\omega t) \)
- \( \mathcal{E}_0 e^{-\kappa z} \exp (i\omega t) \)
- \( \mathcal{E}_0 e^{-\kappa z} \exp (-ikz) \exp (i\omega t) \)

Roman letters (\( E \)) represent time-dependent quantities which vary like \( \exp (i\omega t) \); script letters (\( \mathcal{E} \)) are independent of time. To obtain the complete expression for an attenuated wave the different symbols must be multiplied by the following factors:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Multiply by</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E )</td>
<td>( \mathcal{E} ) \exp (i\omega t)</td>
</tr>
<tr>
<td>( E_0 ) \exp (-ikz)</td>
<td>( \mathcal{E}_0 ) \exp (-ikz) \exp (i\omega t)</td>
</tr>
<tr>
<td>( \mathcal{E} e^{-\kappa z} )</td>
<td>( \mathcal{E}_0 e^{-\kappa z} \exp (i\omega t) )</td>
</tr>
<tr>
<td>( \mathcal{E}_0 e^{-\kappa z} \exp (-ikz) )</td>
<td>( \mathcal{E}_0 e^{-\kappa z} \exp (-ikz) \exp (i\omega t) )</td>
</tr>
</tbody>
</table>

The following symbols will be frequently used:

- \( c \) = free-space velocity of electro-magnetic waves.
- \( e \) = charge on an electron (when numerical values are inserted this will be negative).
- \( H_0 \) = magnitude of the imposed magnetic field \( \mathbf{H}_0 \).
- \( k \) = angular wave number = \( 2\pi/\lambda \).
- \( l = 0, \) or \( 1/8, \) for Sellmeyer or Lorentz theory respectively (see Chapter 15).
- \( m \) = mass of electron.
- \( n \) = complex refractive index = \( \mu - i\chi \).
- \( N \) = number density of electrons.
- \( x, y, z \) = position coordinates.
- \( E \) = electric wave-field.
- \( \mathbf{H} \) = magnetic wave-field.
- \( \mathbf{P} \) = volume polarization of the medium.

\( \omega \) = angular wave-frequency.
\( \mu \) = refractive index = real part of \( n \).
\( \kappa \) = absorption coefficient.
\( \chi \) = \( \kappa c/\omega \) = absorption index = negative imaginary part of \( n \).
\( \mu_0 \) = magnetic permittivity of free space.
\( \varepsilon_0 \) = electric permittivity of free space.
\( \theta \) = angle between \( \mathbf{H}_0 \) and Oz.
\( \nu \) = frequency of collision of electrons with heavy particles.

The following special symbols are similar to, but not quite the same as, those introduced by Appleton (1932). They have been recommended by the Union Radio Scientifique Internationale (U.R.S.I.) for use in magneto-ionic theory. It should be especially noticed that the symbols \( X, Y, Z \) replace Appleton's \( x, y, z \) and do not mean the same as his \( X, Y, Z \):

\[ \omega_K = 4\pi Ne^2/\varepsilon_0 m \quad \omega_K^2 \text{ is often used in other works}. \]
\[ \omega_H = \mu_0 H_0 |e|/m. \]
\[ \omega_L = (\mu_0 H_0 e/m) \cos \theta. \]
\[ \omega_T = (\mu_0 H_0 e/m) \sin \theta. \]
\[ X = \omega_K^2/\omega^2. \]
\[ Y = \omega_H/\omega. \]
\[ Y_L = \omega_L/\omega. \]
\[ Y_T = \omega_T/\omega. \]
\[ Z = \nu/\omega. \]
\[ \omega_c = (\mu_0 H_0 e/m)(1/2 \sin^2 \theta/\cos \theta) \]

(This symbol is not included in the recommendations of U.R.S.I.)

It should be noted that \( \omega_H \) and \( Y \) are defined so that their signs do not depend on the sign of \( e \), or the direction of the imposed field, but that the signs of \( \omega_L, \omega_T, \omega_c, Y_L \) and \( Y_T \) depend both on the sign of \( e \) and on whether \( \theta \) is greater than or less than \( \frac{1}{2} \pi \).

Unrationalised units will be used.

In all expressions the electric and magnetic permittivities (\( \varepsilon_0, \mu_0 \)) of free space will be retained. It is then a simple matter to evaluate the expressions in terms of any particular set of units by substituting the appropriate numbers for \( \varepsilon_0 \) and \( \mu_0 \).
PART I. THE DERIVATION OF THE EQUATIONS

CHAPTER 2

THE MACROSCOPIC APPROACH. DERIVATION OF APPLETON'S EQUATIONS BY MEANS OF AN EQUIVALENT CONTINUOUS MEDIUM

2.1. The equivalent continuous magneto-ionic medium

We shall define a magneto-ionic medium as one in which free electrons and heavy positive ions are situated in a uniform magnetic field and are distributed with statistical uniformity, so that there is no resultant space-charge. Sometimes we shall suppose that neutral molecules, with which the electrons can make collisions, are also present.

The magneto-ionic theory is concerned with the propagation of electro-magnetic waves through a medium of this kind. The results will be expressed by Appleton's equations (2.6.6–2.6.10) which it is the purpose of this chapter to develop.

Nearly all the problems considered in this book are of a kind where linear superposition is applicable so that a problem involving non-sinusoidal wave-fields can be dealt with by the method of Fourier synthesis. Wave-fields will therefore be assumed to be sinusoidal and will be represented as complex quantities with the notation explained in §1.5.

In this chapter and in Chapter 4 it will be supposed that the number of electrons in one cubic wave-length is so great that they are equivalent to a continuous distribution of charge, and their movement is equivalent to a continuous distribution of current. The theory developed in this way will be called a macroscopic theory. In Chapters 3 and 5 the theory will be developed differently by considering the detailed effects of the radiation emitted by the individual electrons, and that will be called a microscopic theory.

In considering the effect of a continuous distribution of charge on an electro-magnetic wave an important quantity is the polarisation $P$ defined as the induced dipole moment per unit volume of the medium. This quantity can be related to the movements of the electrons as follows.

The movement of an electron from a point $A$ (fig. 2.1) to a point $B$ is equivalent to leaving the original electron undisturbed at $A$ and adding a dipole as shown in the figure so that it cancels the original charge at $A$. If there are $N$ electrons per unit volume, and if all move through equal distances $r$, the equivalent dipole moment per unit volume, which is the polarisation $P$, is given by

$$ P = N\mathbf{r}. \quad (2.1.1) $$

The movement of the electrons, and hence the polarisation $P$, is produced by the electric field $E$ of the wave, and if there is no applied magnetic field $P$ and $E$ are parallel and the medium is isotropic. The displacement $r$, and the polarisation $P$, are then both proportional to $E$, and

$$ P = \sigma E, \quad (2.1.2) $$

where $\sigma$ is a scalar quantity called the polarisability of the medium.

In the presence of a steady magnetic field, however, the electrons do not necessarily move along the direction of $E$, so that $E$ and $P$ are not necessarily parallel, and the medium is non-isotropic. The isotropic case, where there is no magnetic field, will be considered first.

Suppose that a sinusoidal electro-magnetic wave is travelling through the medium with electric field $E = \mathbf{E}_0 \exp(\mathbf{i} \omega t)$. Suppose, also, that the movements of the heavy positive ions may be neglected, and, for the present, that the electrons do not collide with other particles. The equation of motion of a single electron is then

$$ m\ddot{\mathbf{r}} = \mathbf{eE}, \quad (2.1.3) $$

provided the movements of the neighbouring electrons do not influence its motion. If the motion of one electron were influenced by

* This condition will be removed in Chapter 16.
the motion of the other electrons it would be necessary to include a ‘Lorentz Term’ in the equations as discussed separately in Chapter 15. It is, however, here supposed, in agreement with present ideas, that there is no Lorentz Term and that equation (2.1.3) is correct. This leads to
\[ r = (-e/mw^2)E, \] (2.1.4)
and from (2.1.1)
\[ P = (-Ne^2/mw^2)E. \] (2.1.5)

A comparison of (2.1.2) and (2.1.5) shows that the polarisability \( (\sigma) \) of the medium is given by
\[ \sigma = -Ne^2/mw^2. \] (2.1.6)

The equations (2.1.2) and (2.1.6) describing the behaviour of the equivalent continuous medium will be called the constitutive relations of the medium. They are particularly simple in the present example, where collisions and imposed magnetic field are supposed absent; in later sections they will become more complicated, but whatever the complications, as soon as they have been written down it is no longer necessary to consider individual electrons and the subsequent calculations can be made in terms of the equivalent continuous medium defined by these constitutive relations.

The polarisability \( (\sigma) \) may now be related to the dielectric constant \( (\varepsilon) \) of the medium as follows. The displacement \( D \) is defined by the expression
\[ D = \varepsilon_0 E + 4\pi P, \] (2.1.7)
and in the present case, where \( E \) and \( P \) are parallel this can be written
\[ D = \varepsilon_0 E, \]
where
\[ \varepsilon = 1 + (4\pi/e_0)(|P|/|E|) = 1 + 4\pi\sigma/e_0 \] (2.1.8)
is defined as the equivalent dielectric constant of the medium. It is next necessary to consider the passage of a wave through a continuous medium with this equivalent dielectric constant.

2.2. A wave in a continuous medium of dielectric constant \( \varepsilon \)

Maxwell’s equations for free space can be written
\[ \begin{align*}
\text{curl} H &= \varepsilon_0 \frac{\partial \varepsilon}{\partial t} \hat{E}, \\
\text{div} H &= 0, \\
\text{curl} E &= -\mu_0 \frac{\partial \varepsilon}{\partial t} \hat{H}, \\
\text{div} E &= 0,
\end{align*} \] (2.2.1)
and are all well known to lead to a velocity \( c = 1/\sqrt{(\mu_0 e_0)} \) for waves of all frequencies.

In a medium of dielectric constant \( \varepsilon \) the equations become
\[ \begin{align*}
\text{curl} H &= \varepsilon_0 \frac{\partial \varepsilon (\partial H)}{\partial t}, \\
\text{div} H &= 0, \\
\text{curl} E &= -\mu_0 \frac{\partial \varepsilon}{\partial t} \hat{H}, \\
\text{div} E &= 0,
\end{align*} \] (2.2.2)
and lead to a velocity
\[ v = 1/\sqrt{(\mu_0 e_0)} = c/\sqrt{\varepsilon}. \]

If this velocity is expressed in terms of a refractive index \( n \), so that
\[ v = c/n, \] (2.2.3)
then
\[ n^2 = \varepsilon. \] (2.2.4)

Substitution from (2.1.8) then gives
\[ n^2 = 1 + 4\pi\sigma/e_0, \] (2.2.5)
which is a useful general expression relating the refractive index \( n \) to the polarisability \( \sigma \) of the medium.

If a particular medium is to be considered \( \sigma \) must first be deduced from the constitutive relations, and then inserted in (2.2.5) to give the refractive index. For the medium consisting of free electrons, considered in §2.1, the constitutive relation is given by (2.1.6), and substitution into (2.2.5) gives
\[ n^2 = 1 - 4\pi Ne^2/e_0 mw^2. \] (2.2.6)

The refractive index is less than unity, and reference to (2.1.4), (2.1.5) and (2.1.6) shows that this is because a free charge, of either sign, subjected to an alternating field, is displaced by an electric field in such a direction that it is equivalent to an oscillating dipole in antiphase with the driving field, so that \( \sigma \), in (2.1.6), is negative.
2.3. Waves in a medium which is electrically non-isotropic.

Characteristic waves

Suppose now that the medium is electrically non-isotropic so that the volume-polarisation \( \mathbf{P} \) is not necessarily in the direction of the electric field \( \mathbf{E} \). Consider also a wave which is, in general, not plane-polarised, so that its electric and magnetic field may have components \( E_x, E_y, E_z, H_y, H_z \) along all three coordinate directions. The wave-polarisation* will be described in terms of the complex ratios \( E_y/E_z, E_z/E_x, H_y/H_z \) and \( H_z/H_y \) and it will be shown later that, in general, this polarisation will alter as the wave travels. First, however, an attempt will be made to discover whether there are any waves which can be propagated along any given direction without their polarisations changing. If there are, they will be called ‘characteristic waves’.

Suppose that, in a plane wave travelling with its wave-normal along \( Oz \), all the components of the wave-fields vary with \( z \) like \( \exp(-ikz) \), and do not vary in the directions \( Ox \) and \( Oy \) which lie in the wave-front. The ratios of the different components are then the same for all values of \( z \) so that the wave-polarisation does not change as the wave travels, and it is a ‘characteristic wave’. For this wave two of Maxwell’s equations can be transformed as follows:

The equation \( \text{curl} \mathbf{H} = \mathbf{D} = \varepsilon_0 \mathbf{E} + 4\pi \mathbf{P} \) gives:

\[
+ i k E_y = i \omega (\varepsilon_0 E_x + 4\pi P_x),
\]

\[
- i k E_z = i \omega (\varepsilon_0 E_y + 4\pi P_y),
\]

\[
o = - i \omega (\varepsilon_0 E_z + 4\pi P_z) = i \omega D_z.
\]

* The word ‘polarisation’ is used in two senses to mean (a) the dipole moment per unit volume of the medium and (b) the state of polarisation of the wave. When there is any possibility of confusion we shall refer to these two magnitudes as ‘the volume-polarisation’ and ‘the wave-polarisation’ respectively.

It is interesting to note that the two meanings, now thought of quite differently, originated in the same idea, that of producing opposite polarity at opposite ends of a line. The relation to the volume-polarisation is obvious. In relation to wave-polarisation the term arose when it was thought that light was corpuscular in nature, and to explain the facts of double refraction it was suggested that the corpuscles acquired a polarity along a certain direction, known as the direction of polarisation. The equation \( \text{curl} \mathbf{E} = -\mu_0 \mathbf{H} \) gives

\[
+i k E_y = -i \omega \mu_0 H_x,
\]

\[
- i k E_z = -i \omega \mu_0 H_y,
\]

\[
o = -i \omega \mu_0 H_z.
\]

First, we note from (2.3.6) that \( H_z = 0 \), i.e., the magnetic field of the wave lies entirely in the wave-front. Next from (2.3.3) we see that \( E_z \) is not necessarily zero, so that the electric field of the wave may have a component perpendicular to the wave-front, i.e. in the direction of the wave-normal. This difference between the nature of the magnetic and electric fields arises because it has been assumed that the medium has electric properties but no magnetic properties, so that there is no induced magnetic polarisation corresponding to the electric polarisation \( \mathbf{P} \). It is interesting to note that (2.3.3) shows \( D_z \) to be zero so that \( \mathbf{D} \) lies entirely in the wave-front, although \( \mathbf{E} \) may have a component perpendicular to the wave-front.

Now combine (2.3.1) with (2.3.5) and (2.3.2) with (2.3.4) to give

\[
k^2/\epsilon_0 \mu_0 \omega^2 = 1 + (4\pi/\epsilon_0) (P_z/E_z).
\]

and

\[
k^2/\epsilon_0 \mu_0 \omega^2 = 1 + (4\pi/\epsilon_0) (P_y/E_y).
\]

Since the left-hand sides of these equations are equal, and constant, we see that \( P_z \) and \( P_y \) must vary with \( z \) like \( E_z \) and \( E_y \), and that

\[
P_z/E_z = P_y/E_y.
\]

Only those waves which satisfy (2.3.9) will travel without changing their wave-polarisations. These are the characteristic waves capable of travelling along the \( Oz \)-direction.

Next we use (2.3.7) to deduce the velocity \( v \) of the characteristic wave. Since \( v = \omega/k \) we have

\[
1/v^2 = (\mu_0 \epsilon_0) \{1 + (4\pi/\epsilon_0) (P_z/E_z)\}.
\]

Then, writing \( v = c/n \), where \( c = (\mu_0 \epsilon_0)^{-1/2} \) and \( n \) is the refractive index, we arrive at the expression

\[
n^2 = 1 + (4\pi/\epsilon_0) (P_z/E_z),
\]

or, starting from (2.3.8) instead,

\[
n^2 = 1 + (4\pi/\epsilon_0) (P_y/E_y).
\]
Finally we use (2.3.4) and (2.3.5) to show that

\[ E_y = -\left(\mu_0 \omega/k\right) H_x \quad \text{and} \quad E_x = \left(\mu_0 \omega/k\right) H_y, \quad (2.3.13) \]

so that

\[ E_x/E_y = -H_y/H_x. \quad (2.3.14) \]

If the ratios in (2.3.14) are complex the vector which represents the projection of \( E \) on the wave-front is elliptically polarised, as also is the field-vector \( H \), and (2.3.14) shows that the \( H \)-ellipse and the ellipse which represents the projection of \( E \) on the wave-front are similar and are described in the same sense, but have their major axes at right angles. If \( k \) is real (2.3.13) shows that \( H \) and the projection of \( E \) on the wave-front are at right angles to each other throughout the cycle, but that, if \( k \) is complex, i.e. if there is absorption, the instantaneous magnetic field is not perpendicular to the instantaneous component of the electric field in the wave-front.

Those results of this section which are stated in italics, and those represented by (2.3.11) and (2.3.12), hold quite generally for the propagation of a characteristic wave through a medium with electrical properties even if it is non-isotropic. They hold, for example, for waves propagated through crystalline media. In order to investigate any particular medium it is now necessary to consider in detail how the polarisation \( P \) is related to the electric field \( E \) through the constitutive relations. From these relations the wave-polarisation which will make \( E_x/E_y = P_x/P_y \), as in (2.3.9), is found. This determines the wave-polarisation of the characteristic wave. The magnitude of \( P_x/P_y \) is then deduced, for this particular wave-polarisation, and is inserted in (2.3.11) to give the refractive index of the wave. These calculations are made in the next section for the particular case of a magneto-ionic medium.

2.4. The constitutive relations for a magneto-ionic medium

Suppose that a magneto-ionic medium consists of electrons, with number-density \( N \), situated amongst a neutralising distribution of stationary positive charge in the presence of a steady magnetic field \( H_0 \). Assume that \( H_0 \) lies in the \( xy \)-plane, and has components \( H_L \) along \( OZ \) and \( H_T \) along \( OY \) as in fig. 2.2. Assume also that a sinusoidal wave is passing through the medium along the positive \( OZ \)-direction and that the components of electric field are represented by complex quantities \( E_x, E_y, E_z \) of the form \( E = \varepsilon \exp (i\omega t) \). In the presence of the imposed magnetic field an electric field moves the electrons, in general, not only along its own direction, but also perpendicular to that direction. The equations of motion of an electron are

\[
\begin{align*}
e E_x &= m\ddot{x} - \mu_0 e H_L \dot{y} + \mu_0 e H_T \dot{z}, \\
e E_y &= m\ddot{y} + \mu_0 e H_L \dot{x}, \\
e E_z &= m\ddot{z} - \mu_0 e H_T \dot{x}.
\end{align*}
\]

(2.4.1)

![Fig. 2.2. Coordinate axes.](image)

These equations will now be expressed in terms of the polarisation \( P = \varepsilon \exp (i\omega t) \) by using the relation \( P = N \varepsilon r \), so that, after a little rearrangement, they become

\[
\begin{align*}
\frac{\varepsilon_0}{4\pi} E_x &= -\frac{\varepsilon_0 m \omega^2}{4\pi N e^2} P_x - \frac{i\omega e_0 m}{4\pi N e^2} \left(\frac{\mu_0 H_L e}{m}\right) P_y + \frac{i\omega e_0 m}{4\pi N e^2} \left(\frac{\mu_0 H_T e}{m}\right) P_z, \\
\frac{\varepsilon_0}{4\pi} E_y &= -\frac{\varepsilon_0 m \omega^2}{4\pi N e^2} P_y + \frac{i\omega e_0 m}{4\pi N e^2} \left(\frac{\mu_0 H_L e}{m}\right) P_x, \\
\frac{\varepsilon_0}{4\pi} E_z &= -\frac{\varepsilon_0 m \omega^2}{4\pi N e^2} P_z - \frac{i\omega e_0 m}{4\pi N e^2} \left(\frac{\mu_0 H_T e}{m}\right) P_x.
\end{align*}
\]

(2.4.2)
If these equations are written in terms of the symbols introduced in §1.5 they take the form

\[ (\varepsilon_0 X/4\pi) E_x = -P_x + iY_P P_z - iY_L P_y, \]  
\[ (\varepsilon_0 X/4\pi) E_y = -P_y + iY_P P_z, \]  
\[ (\varepsilon_0 X/4\pi) E_z = -P_z - iY_P P_x. \]  

These constitutive relations show how the components of the polarisation \( P_x, P_y, P_z \) are related to the fields \( E_x, E_y, E_z \), and if there is no imposed magnetic field they reduce to

\[ (\varepsilon_0 X/4\pi) E = -P, \]  

which is the same as (2.1.5) appropriate to this simple case.

It is clear from the more complicated relations (2.4.3), (2.4.4) and (2.4.5) that, when the magnetic field is imposed, the resulting polarisation is not always in the same direction as the applied field. One way of stating this result is to say that the polarisability \( \sigma \) which relates \( E \) to \( P \) is a tensor quantity, and the most concise form of the calculation makes use of this fact. That calculation is presented in Chapter 18, but for the present purpose the theory will be developed by a simple adaptation of the arguments previously used.

### 2.5. The characteristic waves in the magneto-ionic medium

The condition which must be satisfied in a characteristic wave travelling in a medium with a non-isotropic polarisability has already been shown to be (2.3.9)

\[ P_x/E_x = P_y/E_y. \]  

This condition, with the constitutive relations of (2.4.3), (2.4.4) and (2.4.5), these conditions can be elaborated as follows. Substitute into (2.4.5) from (2.3.3) and obtain

\[ (1 - X)P_z = -iY_P P_x. \]  

Now substitute for \( P_z \) in (2.4.3) to obtain

\[ (\varepsilon_0 X/4\pi) E_x = -(1 - Y_P^2/(1 - X)) P_z - iY_L P_y, \]  

or

\[ (\varepsilon_0 X/4\pi)(E_x/P_x) = -(1 - Y_P^2/(1 - X)) - iY_L(P_y/P_x), \]

and rearrange (2.4.4) to give

\[ (\varepsilon_0 X/4\pi)(E_y/P_y) = -1 + iY_P(P_z/P_y). \]  

Now use (2.5.4) and (2.5.5) to express the condition (2.5.1) and obtain

\[ 1 - Y_P^2/(1 - X) + iY_P(P_z/P_y) = 1 - iY_L(P_z/P_y) \]

or

\[ iY_L(P_z/P_y) - Y_P^2/(1 - X) = Y_L(P_z/P_y) + iY_L = 0. \]  

This quadratic equation yields two values for the ratio \( P_z/P_y \), and these correspond to two possible characteristic waves in the medium.

We now define a quantity \( R \), which we will call the 'polarisation of the wave-field', by the relation

\[ R = -H_y/H_x. \]

If \( R \) is a real quantity the magnetic field of the wave is linearly polarised; if \( R \) is complex it is elliptically polarised. We have chosen to represent the wave polarisation in terms of \( H \) rather than \( E \) because \( H \) lies entirely in the wave-front (see §2.3), whereas there is a component \( E_x \) of \( E \) perpendicular to the wave-front. Equations (2.3.9) and (2.3.14) show that

\[ R = -H_y/H_x = E_x/E_y = P_z/P_y. \]

We now derive an expression for \( R (= P_z/P_y) \) by solving (2.5.6) to give

\[ (P_z/P_y) = R = -i \left( \frac{Y_L}{2(1 - X)} \right)^\dagger \left( \frac{Y_T}{4(1 - X)^2 + Y_L^2} \right)^\dagger. \]  

Since \( H_x/H_y = -R \) and \( H_x = 0 \) this expression gives full information about the polarisations of the magnetic wave-fields in the characteristic waves. The two possible waves, corresponding to the upper and lower signs in (2.5.8), have values \( R_u \) and \( R_l \) of \( R \) such that

\[ R_u R_l = 1. \]

The significance of this result and the other properties of the quantity \( R \) will be discussed in Chapter 7.

The quantity \( R \) also gives the ratio \( E_z/E_y \) of those components of the electric wave-field which lie in the wave-front, but in order

* The alternative signs in front of the square root are taken this way round so that (2.5.14) will take its usual form.
to complete the description of the electric wave-field it is necessary also to introduce the longitudinal component $E_x$. This is possible through (2.3.3), (2.5.2) and (2.3.11), which are repeated here.

\[
(\varepsilon_0/4\pi) E_x = -P_x, \quad (2.3.3)
\]
\[
(1 - X) p_x = -i Y_T P_x, \quad (2.5.2)
\]
\[
n^2 = 1 + (4\pi/\varepsilon_0) (P_y/E_y). \quad (2.3.11)
\]

From these equations the quantity $S$, defined to be equal to $E_x/E_z$, is given by

\[
S = E_x/E_z = i Y_T (n^2 - 1)/(1 - X), \quad (2.5.10)
\]

and it can also be shown that

\[
P_x/P_z = -S/(n^2 - 1) = -i Y_T/(1 - X). \quad (2.5.11)
\]

The two characteristic waves, with the two wave-polarisations given by (2.5.8), are propagated with refractive indices given by (2.3.12)

\[
n^2 = 1 - X/(1 - iY_L R), \quad (2.5.12)
\]

Substitution from (2.5.5) into (2.5.12) gives

\[
n^2 = 1 - X/(1 - iY_L R), \quad (2.5.13)
\]

or, after substitution for $R$ from (2.5.8),

\[
n^2 = 1 - X/(1 - iY_L R) \pm [i Y_L/(1 - X)]^2. \quad (2.5.14)
\]

Before discussing the implications of this equation for the refractive index it will be generalised to include the effects of collisions between the electrons and heavy particles.

2.6. The effects of collisions

So far there has been no mention of the collisions which the electrons make with heavy particles such as neutral atoms or positive ions. In Chapter 4 it will be shown that their averaged effect is to modify the equation of motion of a single electron by the introduction of a ‘viscous damping force’ proportional to the velocity $\mathbf{f}$ of the electron, and given by $-mv\mathbf{f}$, where $v$ is the average frequency of collision between an electron and heavy particles. In the absence of an imposed magnetic field the equation of motion of an electron is therefore

\[
m\ddot{\mathbf{r}} + m_\mathbf{v} = e\mathbf{E}, \quad (2.6.1)
\]

instead of

\[
m\ddot{\mathbf{r}} = e\mathbf{E}. \quad (2.6.2)
\]

If the procedure of (2.1.3), (2.1.4) and (2.1.5) is now followed it is found that

\[
\mathbf{P} = [-Ne^2/(m(\omega^2 - i\nu))] \mathbf{E}, \quad (2.6.3)
\]

or, in the nomenclature of § 1.5 and equation (2.4.6),

\[
(e_0 X/4\pi) \mathbf{E} = -(1 - iZ) \mathbf{P}. \quad (2.6.4)
\]

When it is required to investigate the wave which passes through a magneto-ionic medium in which the electrons make collisions with heavy particles the first step is to introduce terms $mv\mathbf{f}, mv\mathbf{f}$, $mv\mathbf{f}$ on the right-hand side of the equations of motion (2.4.1) of the electrons. The equations are then dealt with exactly as in § 2.4, and, instead of (2.4.3), (2.4.4) and (2.4.5) there are obtained

\[
\begin{aligned}
(e_0 X/4\pi) E_x &= -(1 - iZ) P_x + i Y_T P_z - i Y_L P_y, \\
(e_0 X/4\pi) E_y &= -(1 - iZ) P_y + i Y_L P_x, \\
(e_0 X/4\pi) E_z &= -(1 - iZ) P_z - i Y_T P_x.
\end{aligned} \quad (2.6.5)
\]

From these constitutive relations the properties of the characteristic waves can be deduced by the same procedure as before. It is then found that the equations with the numbers given on the left below take the following modified forms:

\[
R = -H_y/H_x = E_y/E_x = P_y/P_x, \quad (2.6.6)
\]
\[
R = -iY_L/(1 - X - iZ) \mp \left[ i Y_L/(1 - X - iZ)^2 + Y_L^2 \right]^{1/2}, \quad (2.6.7)
\]
\[
S = E_x/E_y = i Y_T (n^2 - 1)/(1 - X - iZ), \quad (2.6.8)
\]
\[
P_x/P_z = -S/(n^2 - 1) = -i Y_T/(1 - X - iZ), \quad (2.6.9)
\]
\[
n^2 = 1 - iZ - i Y_L/(1 - X - iZ) \pm [i Y_L/(1 - X - iZ)^2 + Y_L^2]^{1/2}, \quad (2.6.10)
\]

In this more general case, (2.5.13) is replaced by

\[
n^2 = 1 - X/(1 - iZ - i Y_L R). \quad (2.6.11)
\]
so that, in general, \( n \) is complex, and the waves are attenuated as a result of the collisions made by the electrons.

The relation 
\[
R_u R_t = 1 \tag{2.6.12}
\]

still holds as in the absence of collisions (2.5.9).

The equations (2.6.6)-(2.6.10) will be called Appleton's equations or the equations of the magneto-ionic theory; it is the main purpose of this book to discuss in detail several of the important deductions which can be made from them. They represent the behaviour of the two characteristic plane waves which can be propagated along the \( Ox \)-direction in a homogeneous magneto-ionic medium with their polarizations unchanged. If a wave starts to travel in the direction \( Ox \) with a polarisation different from one of these it will be split into these two waves, with relative amplitudes determined by the original polarisation. Each wave will then travel with its own characteristic velocity and attenuation, and as the combined wave progresses the two components will add to give a resulting polarisation which changes as the wave travels. Only if the original wave has one of the 'characteristic' polarisations will it travel with its polarisation unchanged.

The velocities, the absorptions, and the wave polarisations of the two waves depend amongst other things on the quantities \( Y_L \) and \( Y_R \) and hence on the angle (\( \theta \)) between the direction of propagation and the direction of the imposed magnetic field \( H_0 \).

**CHAPTER 3**

**THE MICROSCOPIC APPROACH. THE PHYSICAL REASON FOR THE CHARACTERISTIC WAVES AND THEIR DISPERSION**

3.1. The microscopic explanation of dispersion

In the previous chapter the dispersion of waves travelling in a magneto-ionic medium was considered, and it was shown that two characteristic waves could be propagated in any given direction, each with its characteristic polarisation. The method of calculation was a 'macroscopic' one in which the detailed movements of the electrons were averaged out so as to constitute a volume polarisation in a continuous medium to which Maxwell's equations could be applied.

In this chapter the same phenomenon is discussed from a different point of view. Attention will be directed to the elementary wavelets which are scattered by the individual oscillating electrons, and it will be shown how these add to the main wave to produce a modified wave which travels with a different velocity. A theory of this kind will be called a microscopic theory. The development of the full expression for the magneto-ionic medium is naturally somewhat complicated, so the argument will be developed in steps.

As an introduction the case of an isotropic medium will be considered in general terms, and at first (§§3.2, 3.3, 3.4) it will be assumed that the waves re-radiated by the oscillating electrons are weak compared with the main wave. Next a form of calculation will be used (§§3.5, 3.6, 3.7) suited for dealing with the case where the re-radiated wave may have considerable intensity, and it will be shown that the result (2.2.6) deduced from the macroscopic theory is arrived at. Finally, the case of the magneto-ionic medium will be discussed, the reason for the special properties of characteristic waves will be explained, and their polarisations and refractive indices will be deduced.
3.2. The wave scattered from an element of volume

The dispersion of waves has been considered in terms of a microscopic theory by Darwin (1924) and Hartree (1929, 1931). Darwin considered waves incident on a slab of non-isotropic material and deduced the nature of the transmitted and reflected waves. Hartree discussed the propagation of waves through an assembly of free electrons, as in the ionosphere, but he included the effect of the Lorentz term which is now thought to be unnecessary. In his first paper he considered a medium with no imposed magnetic field, and in the second paper the effect of a field was included. Because these three papers are completely rigorous they involve considerable mathematical complications. For the purposes of this book we present a simpler method in §3.5. The justification of some of the assumptions in this section can be seen from the fuller treatments of Darwin and Hartree.

First consider the scattering from a small element of volume \(dV\) in an isotropic medium. The arguments of §2.1 show that the \(x\)-component of dipole moment induced in the volume \(dV\) will be given by

\[
P_{20} \, dV = \sigma E_{20} \, dV,
\]

where \(P_{20}\) is the volume polarisation produced in the medium by a field \(E_{20}\) and \(\sigma\) is the (possibly complex) polarisability. The wave radiated to a distance, \(r\), by the oscillating dipole moment \(P_{20} \, dV\) in a direction making an angle \((\pi/2 - \theta)\) with the line of the oscillation, has an electric field of magnitude \(\delta E\), given by

\[
\delta E = \frac{\omega^2 \sigma}{e_0 c^2} \frac{(\cos \theta)}{r} e^{(-ikr)}.
\]

(3.2.2)

If \(E_{20}\) is a wave field of angular frequency \(\omega\), combination of (3.2.1) and (3.2.2) gives

\[
\delta E = \left(\frac{\omega^2 \sigma}{e_0 c^2} \frac{(\cos \theta)}{r}\right) E_{20} \exp(-ikr) \, dV
\]

\[
= \frac{k^2 \sigma}{e_0} \left(\frac{(\cos \theta)}{r}\right) E_{20} \exp(-ikr) \, dV.
\]

(3.2.3)

3.3. Scattering from a thin slab

Suppose now that a plane-polarised plane wave given by

\[
E_x = E_{20} \exp[i(\omega t - kx)] = E_{20} \exp(-ikx)
\]

(3.3.1)

is incident normally on a thin slab of material of thickness \(\Delta z\) and of polarisability \(\sigma\) situated at the plane \(z = 0\). To calculate the resultant wave radiated by the whole slab to a point \(Q\) distant \(d\) from the slab (fig. 3.1) we use an argument of the kind employed in the theory of physical optics. The slab is first divided into elementary annular zones by finding its intersection with spheres of radii \(d + \delta, d + 2\delta, d + 3\delta, \ldots, (d + n\delta), \ldots\), drawn with \(Q\) as a centre. If \(\delta\) is small the whole of any one zone may be taken to be

![Fig. 3.1. Division into elementary zones.](image-url)
expected if the slab had a refractive index \( \mu \) different from unity, so that the phase lag \( \Delta \phi \) would be

\[
\Delta \phi = k(\mu - 1)\Delta z. \quad (3.4.2)
\]

If this value of \( \Delta \phi \) is equated to that in \((3.4.1)\) there results the following expression for \( \mu \) in terms of the polarisability of the medium

\[
\mu - 1 = \frac{2\pi\sigma}{\varepsilon_0}. \quad (3.4.3)
\]

Now \((2.2.5)\), derived on the macroscopic theory, shows that

\[
\mu^2 - 1 = \frac{4\pi\sigma}{\varepsilon_0}. \quad (3.4.4)
\]

Equations \((3.4.3)\) and \((3.4.4)\) are seen to be equivalent for the case considered here, where the scattered field \( |\Delta E'| \) is small compared with the main field \( |E_{\infty}| \) so that \( \mu \approx 1 \).

The expressions deduced in this section are valid whatever the variation of \( \sigma \) with frequency. If \( \sigma \) is complex \( E_{\infty} \) and \( \Delta E' \) are not in quadrature, the change of phase is accompanied by a change of amplitude, and \( \mu \) becomes complex, but \((3.4.3)\) still applies. In the case of a medium containing free electrons \( \sigma \) is real and is given by \( \sigma = -Ne^2/m_0\varepsilon_0 \) \((2.1.6)\), and fig. 3.3(b) represents the wave-fields. In this case the wave has its phase advanced in passing through the slab.

3.5. Propagation through a continuous medium when the polarisability may be large. Fundamental concepts

In this and the two following sections a different method of calculation will be used and the microscopic theory will be extended to deal with the propagation of a wave through a continuous uniform medium, even when the polarisability \( \sigma \) is not small. The fundamental concept is as follows. A uniform wave travelling through a medium of this kind can be thought of as composed entirely of the elementary waves scattered backwards and forwards by all the induced volume polarisations in the medium. Thus the wave at the plane \( z = 0 \) is made up of the wave scattered to this plane from all the elementary slabs at planes where \( z \) has other values. The waves scattered from each slab must be assumed to travel from the slab to \( z = 0 \) with the free-space velocity; any departure from this velocity which might be thought to occur arises because of secondary radiation from other planes which the wave passes through, and all these secondary radiations are going to be taken into account; they must not be added in twice.

In this formulation it should be noted that there is nothing corresponding to the original incident wave of \( \S 3.4 \). Here the medium is supposed to be infinite in extent and the solution is obtained by writing down the condition of self-consistency; that if a wave is to travel it will be such that it excites scattered waves which are just sufficient to constitute the original wave at each point.

3.6. Calculation of the self-consistent wave

Suppose that the self-consistent wave mentioned in the previous section is travelling in the medium along the positive \( z \)-axis, with refractive index \( \mu \), so that it is represented by

\[
E = E_0 \exp(-i\mu k z). \quad (3.6.1)
\]

![Fig. 3.4. Scattering to the plane \( z = 0 \) from thin slabs on the positive and negative sides.](image)

Let us write down the resultant of the secondary waves which are scattered from all points in the medium to the plane \( z = 0 \) and then, according to the principle mentioned in the previous paragraph, equate it to the field of the main wave itself. Consider first the wave scattered from a thin slab contained between the planes \( z \) and \( z + \Delta z \) (fig. 3.4). The wave-field \( E \) at this plane induces dipole moments in the slab which sends waves back to the plane \( z = 0 \) with the free-space velocity, and the total field \( \Delta E \) scattered from the whole slab is, from \((3.3.5)\),

\[
\Delta E = -ik(2\pi\sigma/\varepsilon_0)E \exp(-ik|z|)dz. \quad (3.6.2)
\]
It should be noted that $|z|$ occurs in this expression because there is a retardation in the phase of the wave as it travels from the scattering slab to the plane $z = 0$, irrespective of the direction in which it is travelling. It is now convenient to substitute for $E$ from (3.6.1) and to obtain

$$\Delta E = -ik(2\pi\sigma/\varepsilon_0) E_0 \exp\{-ik(|z| + \mu z)\} \, dz. \quad (3.6.3)$$

The total field scattered to $z = 0$ from the slabs situated at all values of $x$, both positive and negative, is now got by integrating this expression from $z = -\infty$ to $z = +\infty$, and this total field is the one that must be equated to $E_0$; hence

$$i = -ik(2\pi\sigma/\varepsilon_0) \int_{-\infty}^{+\infty} \exp\{-ik(|z| + \mu z)\} \, dz. \quad (3.6.4)$$

If the integral is split into two parts, for $z > 0$ and $z < 0$, the modulus sign round $z$ can be dropped, and there results

$$i = -ik(2\pi\sigma/\varepsilon_0) \left[ \int_{0}^{\infty} \exp\{-ik(1 + \mu) z\} \, dz + \int_{-\infty}^{0} \exp\{-ik(1 - \mu) z\} \, dz \right], \quad (3.6.5)$$

or

$$i = -ik(2\pi\sigma/\varepsilon_0) \left[ \frac{1}{ik(1 + \mu)} + \frac{1}{ik(1 - \mu)} \right]$$

$$= -4\pi\sigma/\varepsilon_0(1 - \mu^2). \quad (3.6.6)$$

Hence

$$\mu^2 - 1 = 4\pi\sigma/\varepsilon_0. \quad (3.6.7)$$

This expression is the same as (2.2.5) derived from a macroscopic theory. It appears, therefore, that the only plane wave which can travel in the homogeneous medium is the self-consistent one in which the field of the wave is the resultant of all the wavelets scattered from the electrons moving under the influence of the wave itself.

* The evaluation of the integrals in (3.6.5) involves difficulties with the limits, since the integrals oscillate according to where, at infinity, they are terminated. These terminations correspond to the two boundaries of the medium at $+\infty$ and $-\infty$, and the oscillations correspond to the standing wave set up at those boundaries. The result given here is the mean value obtained when an average is taken over all positions of the boundaries. The problem is discussed by Darwin (1924).

3.7. Discussion of the self-consistent wave

It is of interest to investigate the relative magnitudes of the total wave-components scattered in the forward and backward directions. Equation (3.6.3) shows that the amplitudes $\Delta E$ of the total waves scattered in the forward and backward directions from elementary slabs of the same thickness are the same. The phases, however, vary with distance in a different way in the two cases. For the wave scattered forward from a plane situated at a negative value of $z$ the phase is proportional to $|z|(1 - \mu)$, while for that scattered backward from a plane with a positive value of $z$ it is proportional to $|z|(1 + \mu)$. The 'space-rate of change' of phase is therefore more rapid for backward than for forward scattering. When the effect of successive elementary slabs is added up by means of the two integrals of (3.6.5) it is clear that backward scattering gives a smaller resultant than forward scattering because the phase changes with distance more rapidly. The resultant waves have amplitudes in the ratio (backward)/(forward) = $(1 - \mu)/(1 + \mu)$. When $\mu = 1$ the wave scattered backward is very much less than that scattered forward, and a good approximation is obtained by considering scattering only in the forward direction. This was done in §3.4.

The analysis presented in this section is not completely rigorous because it has been supposed that only the radiation fields of the oscillating electrons are important. The full calculation, carried out by Hartree (1931 a), using the Hertz Vector, includes the induction and electrostatic fields, and is necessary for a rigid treatment.

3.8. The characteristic waves of the magneto-ionic theory

In this section it will be shown, by considering the scattering from the electrons in a magneto-ionic medium, that a self-consistent wave can exist only if it has one or other of two special polarisations, and these will be found to correspond to those of the two characteristic waves discussed in Chapter 2.

Suppose that a wave of complicated polarisation is capable of travelling in the direction $Oz$ without change of form, and with a velocity corresponding to a refractive index $\mu$, so that it is represented by $E = E_0 \exp(-i\omega z)$. Take axes as shown in fig. 2.2 and resolve the steady imposed magnetic field $H_0$ into components
Suppose the wave has fields $E_x, E_y, E_z$ which act on the electrons to produce polarisations $P_x, P_y, P_z$. The equations of motion of the electrons give rise to the constitutive relations of equations (2.4.3), (2.4.4) and (2.4.5)

$$
\begin{align*}
(e_0X/4\pi)E_x &= -P_x + iY_TP_x - iY_LP_y \\
(e_0X/4\pi)E_y &= -P_y + iY_LP_x \\
(e_0X/4\pi)E_z &= -P_z - iY_TP_y.
\end{align*}
$$

(3.8.1)

The applied fields $E_x, E_y, E_z$ produce a motion of the electrons, described by $P_x, P_y, P_z$ which in general will be elliptical in some skewed plane. The fields arriving at a plane $z$ due to re-radiation from the electron orbits distributed throughout the medium must make up the original wave at the plane $z$. This wave will have the same polarisation everywhere only if, at each point, the resultant of all the re-radiated waves has the same polarisation as the main wave. This condition implies that each radiated wavelet must itself have the same polarisation. If $\delta E_x$ and $\delta E_y$ are the component fields radiated by a single electron $\delta E_x/\delta E_y = P_x/P_y$ and, if the above-mentioned condition is to hold, it follows that

$$
E_x/E_y = \delta E_x/\delta E_y = P_x/P_y.
$$

(3.8.2)

This is the same as the condition (2.3.9) derived from a consideration of Maxwell’s equations. The ‘microscopic’ point of view shows its simplicity and importance.

### Chapter 4

**Absorption on the ‘Macroscopic’ Theory**

#### 4.1. Introduction

In this chapter an account will be given of the absorption which a wave suffers because the electrons collide with heavy particles. First the meaning of a ‘collision’ will be discussed, and then the effect of these collisions on the equation of motion of an electron will be considered. This will lead to the expressions which have already been quoted in the ‘macroscopic’ derivation of the general magneto-ionic equations in §2.6. From the general magneto-ionic equations some useful expressions will be derived for the absorption coefficient.

#### 4.2. The frequency of collision of an electron with heavy particles

Simple gas-kinetic theory shows that if a gas consists of similar atoms or molecules with radii $r_\alpha$, so that a collision will occur if the centres of two of them approach within a distance $2r_\alpha$, then the mean free path $l_\alpha$ is given approximately by

$$
l_\alpha = \frac{1}{\pi(2r_\alpha)^2 n_\alpha}. \quad (4.2.1)
$$

where $n_\alpha$ is the number-density of the atoms or molecules. For the present purpose it is not necessary to consider refinements of this simple theory.

If an electron, of negligible radius, moves amongst the particles of this gas it is supposed that a collision will occur when its centre approaches within a distance $r_\alpha$ of the centre of a particle, so that its mean free path $l_\alpha$ is given approximately by

$$
l_e = \frac{1}{\pi r_\alpha^2 n_\alpha}. \quad (4.2.2)
$$

Unless the electrons, of small radius, are very much more numerous than the heavy particles, collisions amongst themselves may be neglected.
MAGNETO-IONIC THEORY

If the average velocity \( v_e \) of the electrons is known, the average frequency \( (\nu_{ea}) \) of their collisions with neutral particles is determined by the relation

\[
\nu_{ea} = \frac{v_e}{c_e}.
\]

(4.2.3)

In its simplest form gas-kinetic theory gives the velocity \( v_e \) in terms of the temperature \( T \) through the relation

\[
\frac{1}{2}mv_e^2 = \left( \frac{3}{2} \right) kT
\]

or

\[
v_e = (3kT/m)^{1/2}.
\]

(4.2.4)

Combination of (4.2.2), (4.2.3) and (4.2.4) gives

\[
\nu_{ea} = \frac{n_e^2 n_a(3kT/m)^{1/2}}{2}.
\]

(4.2.5)

The magnitude of \( r_a \) is usually assumed to be that appropriate to air at the ground, and is of the order 10^{-8} cm. If this magnitude is inserted in (4.2.5) \( \nu_{ea} \) is found to be given by

\[
\nu_{ea} = 3.7 \times 10^{-8}(T/300)^{3/2} n_a \text{ sec}^{-1},
\]

(4.2.6)

with \( n_a \) expressed in cm^{-3}. If, at a height of 100 km, the magnitudes appropriate to the ionosphere are taken to be \( T = 300^\circ \text{K}, n_a = 10^{13} \text{ cm}^{-3} \), then \( \nu_{ea} \) is calculated to be 3.7 \times 10^{4} \text{ sec}^{-1}.

Suppose next that an electron moves through a gas which consists partly of uncharged atoms or molecules and partly of charged atoms or molecules in the form of ions. Its motion is influenced by an uncharged particle only when it approaches within a distance \( r_a \), and then it undergoes a sudden deflection. But its motion may be considerably influenced by an ion at a much greater distance, because the inverse-square-law forces between charges drop off so slowly. The effective radius for collision between an electron and an ion may, under some conditions, be much greater than its radius for collision with a neutral particle. It is possible to estimate the order of magnitude of this effective radius in the following way. Suppose that the deflection of the electron is sufficiently great for the event to count as a ‘collision’ if it reaches a distance \( r_{eff} \) from the ion such that its potential energy is then equal to its original kinetic energy so that

\[
\frac{1}{2}kT = \frac{e^2}{e_0} r_{eff}.
\]

(4.2.7)

The collision frequency \( \nu_{ei} \), between electrons and ions, may now be derived from (4.2.5) by writing the number-density \( (n_i) \) of ions instead of \( n_a \) and the value of \( r_{eff} \), from (4.2.7), instead of \( r_a \). There then results the expression

\[
\nu_{ei} = \frac{n_i^2 (3kT/m)^{1/2}}{2n_a}
\]

or

\[
\nu_{ei} = (4\pi/3)^{1/2}(e^4/e_0 m^2)(kT)^{-1} n_i.
\]

(4.2.8)

Insertion of numerical values into (4.2.8) yields

\[
\nu_{ei} = 5.1 \times 10^{-4} n_i(T/300)^{-1}.
\]

(4.2.9)

It is noticeable that \( \nu_{ei} \propto T^{-1} \), whereas \( \nu_{ea} \propto T^{1/2} \). The reason for the difference is that, for collisions between electrons and charged ions, the effective cross-section for collisions, given by \( \pi r_{eff}^2 \), depends on the velocity of the electron, and from (4.2.7) it is seen to be proportional to \( T^{-2} \). With the velocity proportional to \( T^{1/2} \) from (4.2.4) it can then be seen that \( \nu_{ei} \) is proportional to \( T^{-1} \).

The argument given here provides no more than an order of magnitude for the collision frequency of electrons with ions, but it leads to an expression of the correct form. A full investigation undertaken by Cowling (1945) leads to the expression

\[
\nu_e = \nu_{ea} + \nu_{ei}
\]

\[
= 1.8 \times 10^{-6}(T/300)^{1/2} n_a + 6.1 \times 10^{-9}(T/300)^{-1} n_i.
\]

(4.2.10)

It will be noticed that each ion contributes as much to the collision frequency as \( f \) neutral particles, where

\[
f = 3.4 \times 10^6(300/T)^{3/2}.
\]

(4.2.11)

As the temperature increases the ions become relatively less important, because, at the great velocities concerned, their collision cross-sections become small. Ions and neutral particles are equally effective at a temperature of about \( 1.8 \times 10^6 \text{ deg. K} \).

It is of interest to consider the application of (4.2.10) to the terrestrial ionosphere. In the \( E \) region it is thought that the temperature is of the order of \( 300^\circ \text{K} \), and \( n_a = 10^{13} \text{ cm}^{-3} \). The collisions with ions will therefore be important only if \( n_i > 10^8 \text{ cm}^{-3} \). Now it is known that the number-density \( (n_i) \) of electrons in this

\* This result implies that the effective cross-section is proportional to \( n_i \), which is the well-known result derived in the analogous calculation for Rutherford scattering of charged particles.
region is about $10^8$ cm$^{-3}$, so that the heavy positive ions corresponding to them could not provide a large enough value of $n_e$ to influence the collision frequency appreciably. If, in addition to the positive ions which accompany the electrons, there were also enough pairs of heavy negative and positive ions of molecular mass they might influence the collision frequency. The negative ions would, however, have to be about 1000 times as numerous as the electrons for this to happen, and there is good evidence (Bates & Massey, 1951), of a different kind, that they are much less numerous. In the $E$ region, therefore, it is supposed that all the important collisions are with neutral molecules.

In the $F$ region the matter is not so simple. The measurements available at present seem to show that $\nu_e = 10^8$ sec$^{-1}$ and $n_e = 10^8$ cm$^{-3}$. If, as is very likely in this region, there are no negative ions, then $n_e = n_i$ and (4.2.10) shows that the measured collision frequency could be explained in terms of collisions with ions alone, if $T \approx 300^\circ$ K. If collisions with neutral particles were to be unimportant this explanation would require that the molecular density in the $F$ region was not greater than about $10^{10}$ cm$^{-3}$, and this would be reasonable if the temperature were as low as $300^\circ$ K. The ionosphere model which leads to this explanation might be called the 'low-temperature model'.

Another model of the $F$ region is, however, often suggested, in which the temperature is supposed to be great, say of the order $3000^\circ$ K. Then, if we write $n_i = n_e = 10^6$ as before, (4.2.10) shows that collisions with heavy ions can contribute a term of magnitude only about $0.2 \times 10^6$ sec$^{-1}$ to the collision frequency. The rest of the observed value of $10^8$ sec$^{-1}$ must therefore come from collisions with neutral atoms and requires that $n_n \approx 10^{14}$ cm$^{-3}$. This atomic density would be available on this 'high-temperature model'.

Two different models of the $F$ region are, therefore, possible at present. In the 'low-temperature model' the gas has a low density, an appreciable fraction of it is ionised, and collisions with ions are of major importance. In the 'high-temperature model' the gas has a high density, only a small fraction is ionised, and collisions with neutral molecules are of major importance. Measurements of collision frequency cannot alone determine which of the two models is the more satisfactory.

In the solar ionosphere there are no neutral particles and the temperature is great. $\nu_{ei}$ is therefore correspondingly small, and, indeed, if the temperature were very great the effect of collisions might become negligible.

It might be thought that the frequencies with which an electron collides with neutral particles or with ions would be seriously altered if a steady magnetic field were imposed on the medium, but full statistical analysis, such as that of Cowling, shows that there is no significant change.

### 4.3. The average effect of collisions on the motion of an electron

In Chapter 2 it was supposed that, apart from the effects of the applied field, the electrons were at rest, and equations were developed to represent their motion. It is now necessary to consider the electrons to be in random gas-kinetic motion, and to be colliding with heavy particles with an average frequency $\nu$. It will be convenient to distinguish between the random gas-kinetic velocity and the ordered oscillatory velocity produced by the action of an imposed wave-motion, and it is necessary to use a statistical argument and to derive an equation for an 'average' electron. The following calculation is based on one first given by Appleton & Chapman (1932).

The equation of motion of a free electron subjected to an alternating electric field $e \exp(i\omega t)$ is

$$m\ddot{x} = e\dot{e} \exp(i\omega t),$$

from which

$$v = \dot{x} = -i(e\dot{e}/m\omega) \exp(i\omega t) + C.$$  \hspace{1cm} (4.3.1)

The electron under consideration has both a random (gas-kinetic) velocity and an ordered (oscillatory) velocity, and it is the ordered velocity which is given by (4.3.1). Just after a collision the ordered velocity must be zero, so that, if the collision took place at a time $t_i$, the integration constant $C$ would be given by

$$C = +i(e\dot{e}/m\omega) \exp(i\omega t_i).$$ \hspace{1cm} (4.3.2)

At a subsequent time $t$, given by

$$t = t_i + \theta,$$ \hspace{1cm} (4.3.3)
MAGNETO-IONIC THEORY

the oscillatory (ordered) velocity \(v\) is, from (4.3.1), given by
\[
v = -i(e\mathcal{E}/m\omega) \{ \exp(i\omega t) - \exp(-i\omega t) \}
\]
\[
= -i(e\mathcal{E}/m\omega) \exp(i\omega t) \{ 1 - \exp(-i\omega \theta) \}
\]
(4.3.4)
for \(t > t_1\). Equation (4.3.4) gives the velocity, at a time \(t\), of an electron which made its last collision at a previous time \(t_1\) such that \(t - t_1 = \theta\). It is now necessary to find how many electrons made their last collision at times within a small range \(\theta\) to \(\theta + d\theta\). If the times between collisions have a Poisson distribution and if the average time is \(1/\mu\), then ordinary gas-kinetic theory shows that, out of \(N\) electrons the number
\[
Nv \exp(-\theta v) d\theta
\]
will, on the average, have made their collisions between times \(\theta\) and \(\theta + d\theta\) earlier. The average oscillatory velocity of all the electrons at the time \(t\) is therefore given by
\[
\bar{v} = \frac{1}{N} \int_0^\infty \left[ -i(e\mathcal{E}/m\omega) \exp(i\omega t) \{ 1 - \exp(-i\omega \theta) \} \right] \{ Nv \exp(-\theta v) \} d\theta
\]
\[
\quad = -i(e\mathcal{E}/m\omega) \exp(i\omega t) \int_0^\infty \left[ \exp(-\theta v) - \exp(-(i\omega + v)\theta) \right] d\theta
\]
\[
\quad = \{e\mathcal{E}/m(i\omega + v)\} \exp(i\omega t).
\]
(4.3.6)

It is possible to write down a fictitious equation of motion, for a single electron, of such a form that it will give the average result of (4.3.6), and then to use it in all subsequent calculations. The equation is
\[
m\ddot{x} + g' \dot{x} = e\mathcal{E} \exp(i\omega t),
\]
(4.3.7)
so that
\[
\dot{x} = \{e\mathcal{E}/(i\omega m + g)\} \exp(i\omega t),
\]
and comparison with equation (4.3.6) shows that
\[
g' = mv.
\]
(4.3.8)

On the average, therefore, the effect of collisions can be allowed for by including the term \(g'\dot{x}\) in (4.3.7) and writing \(g = mv\). The consequences of this modification will be discussed in the next section.

It is interesting to notice that the calculation given here follows closely that given by Lorentz in his Theory of Electrons in Note 56.

ABSORPTION ON 'MACROSCOPIC' THEORY

He was there dealing, however, with bound electrons, for which the appropriate equation was
\[
m\ddot{x} + g' \dot{x} + ax = e\mathcal{E} \exp(i\omega t),
\]
and for this case he found that
\[
g' = 2mv,
\]
a value twice as great as that found here for free electrons. The difference arises because he calculated the average displacement of the electrons whereas here we have calculated their average velocity.

The above calculation, first given by Appleton & Chapman (1932), is approximate only, and several workers (see Bibliography) have tried to make it more precise by considering the detailed statistics of the electrons' collisions. The conclusion appears to be (see Pfister, 1955) that Appleton's magneto-ionic equation provides a good approximation to the refractive index if the quantity \(v\) is taken to be \(\frac{1}{4}\) times the mean collision frequency, except when the collision frequency is large compared with \((\omega + \omega_H)\) for the Ordinary wave, or \((\omega - \omega_H)\) for the Extraordinary. The precise form of the relation under those conditions does not seem to have been decided.

4.4. The absorption coefficient in the magneto-ionic theory

It has been shown in §2.6 that when the equation (4.3.7) of motion of the fictitious equivalent electron is used the complex refractive index \(n\) is given by the expression
\[
n^2 = 1 - \frac{X}{1 - iZ - \frac{1}{4} Y^2/(1 - X - iZ)^2 + \{\frac{1}{4} Y^2/(1 - X - iZ)^2 + \frac{1}{4} X^2\}}
\]
(4.4.1)

To investigate the absorption the complex refractive index is written
\[
n = \mu - i\chi,
\]
(4.4.2)
so that a wave in the medium represented by
\[
E = \mathcal{E}_0 \exp\{i\omega(t - nx/c)\}
\]
can also be written
\[
E = \mathcal{E}_0 \exp\{-\omega\chi x/c\} \exp\{i\omega(t - \mu x/c)\}.
\]
(4.4.3)
This wave is attenuated to \( \exp(-\chi) \) in a distance \( z = c/\omega = \lambda/2\pi \) and travels with a velocity \( c/\mu \). The refractive index is therefore given by \( \mu \) and the absorption coefficient \( \kappa \), defined by

\[
E = iE \exp(-\kappa z),
\]
is given by

\[
\kappa = \omega \chi / c. \tag{4.4-4}
\]

It is next convenient to consider the magnitude of \( \chi \) in the special case when the transmission is along the direction of the earth's magnetic field. This is done by writing \( Y_T = 0 \) in (4.4.1) and \( Y_L \) will be kept as a symbol instead of \( Y \) to emphasise the fact that propagation is along the field. Equation (4.4.1) then gives

\[
(\mu - i\chi)^2 = 1 - X/[1 + |Y_L| - iZ]
\]
or

\[
\mu^2 - \chi^2 - 2i\mu\chi = 1 - \frac{X(1 + |Y_L|)}{(1 + |Y_L|)^2 + Z^2} - \frac{XZ}{(1 + |Y_L|)^2 + Z^2},
\]
from which follows

\[
\mu^2 - \chi^2 = 1 - \frac{X(1 + |Y_L|)}{(1 + |Y_L|)^2 + Z^2}, \tag{4.4-5}
\]
and

\[
2\mu\chi = \frac{XZ}{(1 + |Y_L|)^2 + Z^2}. \tag{4.4-6}
\]

Useful expressions are obtained from (4.4.6) by substituting

\[
\chi = \kappa c / \omega, \quad X = \omega^2 / \omega^2, \quad Y_L = \omega_L / \omega, \quad Z = v / \omega
\]
to obtain

\[
\kappa = \frac{1}{2\epsilon \mu} \left( \omega^2 / |\omega_L|^2 + \nu^2 \right) \tag{4.4-7}
\]
\[
= \frac{1}{2\epsilon \mu} \frac{4\pi Ne^2}{\epsilon_0 m} \left( \omega^2 / |\omega_L|^2 + \nu^2 \right). \tag{4.4-8}
\]

The application of this expression to absorption in a 'non-deviating region' is discussed in §13.6.

Another useful expression is obtained by combining (4.4.5) and (4.4.6) to give

\[
\kappa = \frac{\omega Z}{2\epsilon(1 + |Y_L|)} \left| \frac{1 - \mu^2 + \chi^2}{\mu} \right|
\]
\[
= \frac{\nu}{2\epsilon(1 + |Y_L|)} \left| \frac{1 - \mu^2 + \chi^2}{\mu} \right|. \tag{4.4-9}
\]
CHAPTER 5

THE MICROSCOPIC PICTURE OF ABSORPTION

5.1. Introduction

In this chapter the process of absorption will be considered from the microscopic point of view, and an attempt will be made to see in detail how the energy is removed from the wave by processes involving the collisions of individual electrons. Two different approaches will be used. In the first the energy taken from the wave by the individual electrons will be calculated, and it will be shown that it accounts for the absorption. In the second the wavelets scattered by the individual electrons will be added up, and will be shown to produce the decrease of the field strength which corresponds to absorption. This discussion is an extension of that of Chapter 3, in which it was assumed that the electrons did not make collisions.

5.2. Absorption in terms of energy-transfer

The assumption is first made that the collision frequency is much less than the radio frequency, so that, in a free path between two collisions, the electron makes several oscillations under the influence of the imposed wave as shown in fig. 5.1. In the absence of the wave the free path would be a straight line between $A$ and $B$, but with the wave present there is an oscillatory motion in addition. If the electric field of the wave is along the $x$-direction and is given by $E = \mathcal{E} \cos \omega t$, then the additional oscillatory velocity of the electron is given by $\mathbf{x} = (e\mathcal{E}/m\omega) \sin \omega t$.

![Fig. 5.1. To illustrate the motion of an electron under the influence of a wave-field.](image)

At each instant the ordered energy $U$ in the oscillatory motion, given by $\frac{1}{2}mv^2$, is

$$U = \frac{1}{2}m(e\mathcal{E}/m\omega)^2 \sin^2 \omega t.$$  \hspace{1cm} (5.2.1)

Different electrons make their second collisions (at $B$) at different parts of the cycle, and on the average their ordered energy when they make this collision is given by

$$\overline{U} = \frac{1}{2}m(e\mathcal{E}/m\omega)^2 \sin^2 \omega t = \frac{1}{2}m(e\mathcal{E}/m\omega)^2. \hspace{1cm} (5.2.2)$$

At the second collision ($B$) this ordered energy, which has been abstracted from the wave, is converted into disordered energy corresponding to an increased gas-kinetic velocity of the electron in its next free path. If the gas-kinetic velocity $v_0$ and the ordered velocity $v$, make an angle $\phi$ with each other the total energy is given by

$$\frac{1}{2}mv^2 = \frac{1}{2}m(v_0^2 + v^2 + 2v_0v \cos \phi),$$

and if it is assumed that the gas-kinetic velocities are randomly oriented so that all values of $\phi$ are equally probable, the total energy ($\overline{U_T}$) is, on the average, given by

$$\overline{U_T} = \frac{1}{2}mv_0^2 + \frac{1}{2}mv^2,$$

so that the time averages of the ordered and disordered energies may be added.

It follows that in each free path, on the average, an ordered energy $\overline{U}$ given by (5.2.2), is abstracted from the wave by each electron and is converted into disordered (heat) energy. If each electron makes, on the average, $\nu$ collisions in unit time the power abstracted by it from the wave is

$$\frac{1}{2}m(e\mathcal{E}/m\omega)^2v^2, \hspace{1cm} (5.2.3)$$

Now consider a unit area of the wave-front as it passes a distance $\Delta x$ through a medium containing $N$ electrons per unit volume. There are $N\Delta x$ electrons in the slab so that the total power ($\Delta W$) abstracted is

$$\Delta W = \frac{1}{2}mvN(e\mathcal{E}/m\omega)^2 \Delta x. \hspace{1cm} (5.2.4)$$

Let us now evaluate the power abstracted in this same slab in terms of the attenuation factor $\kappa$, and equate it to that in (5.2.4) to
give a value for \( \kappa \). The time-average power incident on the unit area considered is given by the time-average value of the Poynting vector \( S \), where

\[
S = \frac{EH}{4\pi} = \left( \frac{1}{8\pi} \right) \sqrt{\epsilon_0 \mu_0} n e^{2} \tag{5.2.5}
\]

and \( n \) is the refractive index of the medium. The mean power \( \Delta S \) lost in a thickness \( \Delta x \) is then given by

\[
\Delta S = -(dS/dx) \Delta x = \left( \frac{1}{8\pi} \right) \sqrt{\epsilon_0 \mu_0} n (d\varepsilon^2/dx) \Delta x, \tag{5.2.6}
\]

and if \( \varepsilon = \varepsilon e^{-ax} \) this gives

\[
\Delta S = \left( \frac{1}{8\pi} \right) \sqrt{\epsilon_0 \mu_0} n.2\pi\varepsilon^2 \Delta x. \tag{5.2.7}
\]

If the two expressions (5.2.4) and (5.2.7) for the mean power abstracted are equated there results the following expression for \( \kappa \):

\[
\kappa = \frac{\pi e^2 N \nu}{\epsilon_0 mc \omega^2 n}. \tag{5.2.8}
\]

This may be compared with (4.4.8) derived from the macroscopic theory, which when \( a^2 < \omega^2 \) and \( \omega_L = 0 \) gives a value twice as large. The discrepancy may be ascribed in the following way to the neglect of a ‘transient’ term in the equations of motion of the electron.

It was shown in § 4.3 that the ordered velocity \( (v) \) of an electron which made its last collision at a time \( t_1 \) is given, for \( t > t_1 \), by (4.3.4)

\[
v = -i(e\varepsilon/m\omega)(\exp(i\omega t) - \exp(i\omega t_1)),
\]

or, in real nomenclature,

\[
v = (e\varepsilon/m\omega)(\sin \omega t - \sin \omega t_1). \tag{5.2.9}
\]

The first term in this expression represents the ordered oscillatory velocity so far considered in this section, and the second term represents an additional velocity which remains constant throughout the mean free path. This additional velocity plays a part for free electrons similar to that played by the transient oscillation for bound electrons; it has so far been neglected. It must appear so that, when combined with the initial velocity of the oscillation, the total ordered velocity at the start of the free path shall be zero.

The time-average energy is given by

\[
\frac{1}{2} m(e\varepsilon/m\omega)^2 \left\{ \frac{1}{2} + \sin^2 \omega t_1 \right\}. \tag{5.2.11}
\]

If now a series of electrons are considered, which made their last collisions at all possible times \( t_1 \), the energy averaged for all values of \( t_1 \) is given by

\[
U = \frac{1}{2} m(e\varepsilon/m\omega)^2 \left( \frac{1}{2} + \frac{1}{2} \right) = \frac{1}{2} m(e\varepsilon/m\omega)^2, \tag{5.2.12}
\]

which is twice the value previously obtained in (5.2.2).

Insertion of this new, correct, value for the excess energy gives

\[
\kappa = \frac{2\pi e^2 N \nu}{\epsilon_0 mc \omega^2 n}, \tag{5.2.13}
\]

in agreement with the macroscopic theory.

The microscopic picture of absorption shows, in some detail, how the different factors in the expression (5.2.13) for \( \kappa \) arise. Thus

(a) \( \kappa \propto N \) because each electron abstracts the same power from the wave and converts it to random energy.

(b) \( \kappa \propto \nu \) because the rate of conversion from ordered to disordered energy is proportional to the number of collisions in unit time.

(c) \( \kappa \propto \omega^{-2} \) because the average energy of ordered oscillation of an electron under the influence of a wave is proportional to \( \omega^{-2} \). See (5.2.2).

(d) \( \kappa \propto 1/n(1/\mu) \) because, when the power flux \( S \) is fixed, \( \varepsilon^2 \) is proportional to \( 1/n \) (see (5.2.5)) and the oscillatory energy in the electron is proportional to \( \varepsilon^2 \). It should be especially noted that, in a medium for which \( n < 1 \), the amplitude of the electron’s excursion is greater than when \( n = 1 \), and as \( n \) becomes small the amplitude (and the ordered energy) becomes large.
5.3. Extension to the case where a magnetic field is present (longitudinal case)

It would be unnecessarily complicated to extend the microscopic picture to deal with the case when a magnetic field is superimposed in a general direction; it is, however, profitable to consider the imposition of a field \( \mathbf{H}_0 \) along the direction of propagation. The characteristic wave is then circularly polarised, with one or other sense of rotation, and the electrons move in circles with the same frequency and sense of rotation, and in planes parallel to the wave fronts. For an electron rotating with speed \( v \) round a circular orbit of radius \( r \), under the influence of an electric field \( \mathbf{E} \) of constant strength, rotating with the electron, the equation of motion is

\[
\mathbf{E} e = m v^2 / r \pm \mu_0 H_0 ev.
\]

If \( \omega = v / r \) is the angular frequency of the wave and \( \omega_H = \mu_0 H_0 e / m \) is the angular frequency of free gyration of electrons round the applied field, this may be written

\[
\mathbf{E} e = m o v \pm m o_H v
= m (\omega \pm \omega_H) v.
\]

The velocity is therefore given by

\[ v = \mathbf{E} e / [m (\omega \pm \omega_H)] ; \quad (5.3.1) \]

and the radius \( r \) by

\[ r = \mathbf{E} e / [m o (\omega \pm \omega_H)] , \quad (5.3.2) \]

so that the time-average ordered kinetic energy of rotation is

\[
U_o = \frac{1}{2} m / 2 = \frac{1}{2} m [ \mathbf{E} e / [m (\omega \pm \omega_H)] ]^2 . \quad (5.3.3)
\]

This quantity replaces \( \frac{1}{2} m [\mathbf{E} e / m o]^2 \) of (5.2.2) which is appropriate to the case where there is no applied field. The difference between the factors \( \frac{1}{2} \) and \( \frac{1}{2} \) correspond to the different time-average Poynting vectors appropriate to a plane-polarised wave, and a circularly polarised wave, where the two mutually perpendicular fields produce powers which must be added. Equation (5.3.3) differs from (5.2.2) by the substitution of \( \omega \pm \omega_H \) for \( \omega \), but, apart from that, the ‘microscopic’ calculation proceeds as before, with the result that, if the ‘transient’ velocities are omitted, \( \kappa \) is given by

\[
\kappa = \frac{\pi Ne^2}{\epsilon_0 mc (\omega \pm \omega_H)^2 n} \quad (5.3.4)
\]

instead of by (5.2.8) when \( \omega_H = 0 \).

The ‘transient’ velocities in the presence of a magnetic field are considered in detail in § 14.4, and are shown to correspond to energies equal to those in the ‘forced’ orbits, so that, if they are taken into account as in § 5.2 the final expression for \( \kappa \) becomes

\[
\kappa = \frac{2 \pi Ne^2}{\epsilon_0 mc (\omega \pm \omega_H)^2 n} \quad (5.3.5)
\]

in agreement with the result of macroscopic theory. (See (4.4.8) with \( \omega_H \equiv \omega_L \) and \( (\omega \pm \omega_L)^2 \gg v^2 \).

Equation (5.3.5) shows that the absorption coefficient \( \kappa \) is greater for the extraordinary wave than for the ordinary wave in the ratio \( ((\omega + \omega_H) / (\omega - \omega_H))^2 \). This factor arises because from (5.3.3) it also represents the ratio of the ordered energies in the two cases.

5.4. Absorption in terms of re-radiation from individual electrons

In Chapter 3 it was shown how the refractive index of an electron gas could be explained in terms of the re-radiation from each electron as it was driven by the field of the wave. In a simple calculation concerned with the scattering from a thin slab it was assumed that there were no collisions, and the equation of motion of each electron under the action of the wave was written

\[
\mathbf{m} \ddot{s} = \mathbf{E} e \exp (i \omega t) , \quad (5.4.1)
\]

or, by substituting \( P = N e x \),

\[
\dot{P} = (N e^2 / m) \mathbf{E} e \exp (i \omega t) . \quad (5.4.2)
\]

The elementary field \( \mathbf{E} \) radiated to a distance \( r \) from a small volume \( \mathbf{d}V \) was then written (see (3.2.2))

\[
\mathbf{E} = - (\dot{P} \mathbf{d}V / \epsilon_0 c^2 ) (\cos \theta / r) \exp (- ikr) , \quad (5.4.3)
\]
and it was shown that, when this scattered field was added up, by
the procedure indicated in figs. 3.1 and 3.2, for the whole slab, it
produced at some distance from the slab a field \( \Delta E' \) which was in
quadrature with the main field \( E_0 \) of the wave, as shown in fig.
5.2 (a). When the two fields were added the resultant field led the
original field by a phase angle \( \phi \) which corresponded to an altered
phase velocity.

![Fig. 5.2. Rotating vectors to represent the incident \( (E_0) \) and scattered
\( (\Delta E', \Delta E') \) fields (a) without, and (b) with, absorption.](image)

It is now necessary to introduce the effect of collisions and to
show how they account for absorption of the wave. As in the calculation
of refractive index the wave is considered to pass normally in the
\( z \)-direction through a thin slab of thickness \( \Delta z \) situated at
\( z = 0 \) as in fig. 3.1. The wave scattered from this slab to a point
some distance along \( Oz \) is then evaluated. If, when this scattered
wave is added to the original wave, it is to represent a reduction in amplitude,
corresponding to absorption, there must be another component of the scattered wave,
represented by \( \Delta E'' \) in fig. 5.2 (b), so that, to a first approximation, \( \Delta E'' \)
represents the decrease in amplitude, and the phase angle \( \phi \) is unaltered.

Before it is shown how the scattered field \( \Delta E'' \) arises let us find
how large it must be in order to explain the known absorption. Equation
(4.4.1o) shows that, when there is no imposed magnetic field and \( Y_L = 0 \), the absorption coefficient has the magnitude

\[
\kappa = (\nu/2\mu)(1 - \mu^2),
\]

or, with \( \mu \neq 1 \),

\[
\kappa = (\nu/\mu)(1 - \mu).
\]

The magnitude required for the scattered field \( \Delta E'' \) is now
calculated as follows. In transmission through the distance \( \Delta z \) the
amplitude of the wave is reduced from \( \phi_0 \) to

\[
\phi_0 \exp (-\kappa \Delta z) \Rightarrow \phi_0(1 - \kappa \Delta z)
\]

for small absorption \( \kappa \), and thickness \( \Delta z \). Hence

\[
\Delta E''/E_0 = \kappa \Delta z.
\]

The phase angle \( \phi \) is also given by the following relation (see
(3.4.2)):

\[
\phi = k(1 - \mu) \Delta z = \Delta E'/E_0,
\]

hence

\[
\Delta E'/\Delta E' = \kappa/(k(1 - \mu)).
\]

Substitution from (5.4.5) then yields

\[
\Delta E'' = i(\nu/\omega) \Delta E',
\]

where the factor \( i \) indicates that the phase of \( \Delta E'' \) leads that of \( \Delta E' \)
by \( \frac{1}{4} \pi \), as in fig. 5.2 (b).

In what follows the field re-radiated from a single electron will be
considered. \( \delta E' \) will be taken to represent the re-radiated field
in the absence of collisions, and \( \delta E'' \) the additional field when
collisions occur. It will be shown that

\[
\delta E'' = i(\nu/\omega) \delta E',
\]

so that, after addition of the waves from all the electrons in the slab,
the total scattered fields obey (5.4.6).

Equation (5.4.1) still represents the oscillatory component of
the motion, even when collisions occur, and the magnitude, and
phase, are just as before. The effect of the collisions is merely to
interrupt the smooth oscillatory motion for very short times \( \tau \) which
are negligible compared with the time period of the oscillation, as
shown in fig. 5.3, which indicates the form of the oscillatory motion
with, and without, the collisions. At first sight it seems that the
collisions have little effect on the motion of the electrons. A search
must be made elsewhere for the cause of the extra re-radiated
field \( \delta E'' \).

It will be shown that the extra component of the secondary wave
which is required to explain the absorption arises from the transient
pulses of radiation produced at the collisions. At each collision the
uniform translational velocity of the electron is altered suddenly,
and a small pulse of radiation is emitted. In the absence of a wave
these pulses constitute a 'white' spectrum of radiation whose
energy is in statistical equilibrium with the thermal energies of the
particles. When a wave is present, and an electron has a motion
represented by fig. 5.3, it will be shown that the transient pulses occurring at the collisions are modified. Moreover, this modification represents the addition of a radiation with a single frequency and a definite phase and amplitude, although the collisions occur at random times. It is this coherent component of the radiation which is responsible for the decrease in the amplitude of the wave.

\[
\delta v_x = -\left(\epsilon e \sigma_0 / m_0 \omega\right) \sin \omega t_q
\]  

(5.4.10)

which has been created during the collision, and

\[
v_x' = (\epsilon e \sigma_0 / m_0 \omega) \sin \omega t.
\]  

(5.4.11)

(b) a new uniform velocity of magnitude

\[
v_x = (\epsilon e \sigma_0 / m_0 \omega) \sin \omega t + C.
\]  

(5.4.8)

Just after a collision the additional velocity \(v_x'\) must be zero, just as it would be if the electron started from rest. If, therefore, the collision occurred at a time \(t_q\) the integration constant \(C\) in (5.4.8) is such that, for \(t > t_q\)

\[
v_x = (\epsilon e \sigma_0 / m_0 \omega) (\sin \omega t - \sin \omega t_q).
\]  

(5.4.9)

After collision, therefore, the velocity contains the following components:

(a) a uniform velocity which would be present even in the absence of a wave.
For the special type of field represented in fig. 5.4 equation (5.4.15) leads to
\[ b_n = \frac{2}{T} \sum_q \left[ \sin (2\pi n_q / T) \left\{ \int_{t_q}^{t_q + \tau} \delta \sigma \, dt \right\} \right] \] (5.4.16)
or
\[ b_n = \frac{2}{T} \sum_q [I_q \sin (2\pi n_q / T)]. \] (5.4.17)

Here the integral in (5.4.15) is split up into a sum of the integrals over each of the separate impulses, and the assumption is made that
\[ \delta \sigma(t) \]

represented by the shaded areas, would vary sinusoidally. In this figure it is supposed, for the purposes of illustration, that the shapes of the impulses are all the same, so that their sizes also vary sinusoidally.

the time periods concerned \((T/n)\) are long compared with the duration \(\tau\) of the impulses, so that the term \(\sin (2\pi n / T)\) may be given the constant value \(\sin (2\pi n / T)\) for the impulse occurring at time \(t_q\) and may be taken outside the integral. Substitution from (5.4.13) into (5.4.17) now yields
\[ b_n = \frac{(2/\nu)}{e} \left( e^2/e_0 c^2 m r \right) e_0 \sum_q [\sin (2\pi n_q / T) \sin \omega t_q]. \] (5.4.18)
The summation represents the sum of discrete values of the function \(\sin (2\pi n / T) \sin \omega t\) taken at the instants \(t_q\) when the collisions occur. Since these occur randomly its value approximates to the average value of \(\sin (2\pi n / T) \sin \omega t\) multiplied by the number \((\nu T)\) of collisions occurring in the time \(T\). The average value of \(\sin (2\pi n / T) \sin \omega t\) is zero except when \(2\pi n / T = \omega\), and then it is

- \(T\) can be chosen so that this relation is satisfied for an integral value of \(n\).

\[ \delta\sigma(t) \]

equal to \(\frac{1}{\nu}\), and the \(\Sigma\) factor becomes \(\frac{1}{\nu} \nu T\). Insertion of these values into (5.4.18) shows that
\[ b_n = (e^2 e_0 / e_0 c^2 m r) (\nu / \omega) \] (5.4.19)
when
\[ 2\pi n / T = \omega \]
and \(b_n = 0\) for all other values of \(n\). A corresponding calculation shows that \(a_n\) is zero, because the average value of \(\cos (2\pi n / T) \sin \omega t\) is zero. The expression (5.4.14) for \(\delta\sigma(t)\) can now be written down, but it is more convenient to introduce at once the retardation mentioned just after (5.4.12) and write, by combining (5.4.14) and (5.4.19),
\[ \delta\sigma(t - r/c) = (e^2 e_0 / e_0 c^2 m r) (\nu / \omega) \sin (\omega t - kl). \] (5.4.20)

It thus appears that, so far as the range of frequencies \(\omega / 2\pi \ll 1 / \nu\) is concerned, the field represented in fig. 5.4 corresponds to a single frequency with a definite amplitude and phase, in spite of the fact that the impulses occur at random times. This is indeed obvious if it is remembered that the Fourier analysis of the wave-form is performed by multiplying it by a sinusoidal curve and integrating the result. Only for a curve with the same periodicity as the dotted one will there be a finite result. The single frequency has arisen because the magnitudes of the impulses are related sinusoidally to the times of their occurrence.

It is now necessary to calculate the magnitude of the field \(\delta\sigma\) radiated by the oscillatory component of the electron's oscillation. It is given by
\[ \delta\sigma = -e(x) / e_0 c^2 r = -e^2 e_0 / e_0 c^2 m r \cos (\omega t - kl). \] (5.4.21)
Comparison of (5.4.20) and (5.4.21) shows that, for each electron, the fields \(\delta\sigma\) and \(\delta\sigma'/t\) are such that
\[ \frac{\text{magnitude of } \delta\sigma}{\text{magnitude of } \delta\sigma'} = \frac{\nu}{\omega} \]
and
\[ \text{phase of } \delta\sigma = \text{phase of } \delta\sigma' + \frac{1}{2} \pi, \]
or, if \(\delta\sigma\) and \(\delta\sigma'/t\) are complex quantities representing the fields, then
\[ \delta\sigma' = i(\nu / \omega). \] (5.4.22)

In the previous considerations we were concerned with the fields \(\Delta E\) and \(\Delta E'/t\), radiated by the whole of the slab and arising,
respective, from the oscillatory components, and the randomly timed impulsive components, of the electrons' motions. The fields $\Delta E'$ and $\Delta E''$ are derived by adding the fields $\delta E'$ and $\delta E''$ of the individual electrons by the process indicated in fig. 3.1, but it is not necessary to go into the details of this summation because we have now shown that the complex fields $\delta E'$ and $\delta E''$ arising from the oscillatory and the impulsive components of motion of each electron separately are in the same ratio for all the electrons. Hence the total fields, arising in any way from the summation of these individual fields, are also in the same ratio. Thus

$$\Delta E'/\Delta E'' = \delta E'/\delta E'' = i/(\nu/\omega).$$  \hspace{1cm} (5.4.23)

But this is just the relation (5.4.6) required if $\Delta E''$ is to be responsible for the absorption of the wave.

It has thus been shown that the randomly timed impulses, radiated as shown in fig. 5.4 at the start of the mean free paths in the presence of a wave, constitute just that radiation which is required to reduce the amplitude of the wave by the right amount.

On this picture it appears that the absorption of the wave can be explained in terms of the impulses of radiation associated with the production of the extra uniform velocity at the start of each free path. It might be thought that the impulses of radiation associated with the removal of these extra velocities at the end of each free path should also be taken into account. But there is no relation between the magnitude of the velocity and the time of its destruction, as there is between its magnitude and the time of its creation. The changes of velocity at the ends of the free paths therefore produce quite random fields like those associated with the gas-kinetic velocities. The extra (white) noise-like radiation produced by the destruction of the extra velocity components at the ends of the free paths corresponds to the heating of the electrons by the wave, and represents the extra radiation required to keep the statistical balance between kinetic energy and radiation energy.

### PART II. INTERPRETATION OF THE EQUATIONS

#### CHAPTER 6

**6.1. No imposed magnetic field**

In this chapter the equations of the magneto-ionic theory, developed in previous chapters, will be interpreted in the form of curves on the assumption that there are no collisions between electrons and heavy particles. For this purpose we shall put $Z = 0$ in (2.5.14) and re-write it*

$$n^2 = (\mu - i \chi)^2 = 1 - \frac{X(1 - X)}{(1 - X) - \frac{1}{2} Y^2 \pm \left[ \frac{1}{2} Y^2 + Y^2 (1 - X)^2 \right]^2}. \hspace{1cm} (6.1.1)$$

This expression is always real. When it is positive it is equal to $\mu^2$; when it is negative it is equal to $-\chi^2$. It is convenient for ionospheric theory to consider graphs which show $\mu^2$ and $\chi^2$ as functions of $X (= 4\pi Ne^2/e_0 m^2)$ for different values of $Y (\equiv \omega_H/\omega)$, and of $\theta$, the direction of propagation. $\theta$ enters the equation through $Y_T = Y \sin \theta$ and $Y_L = Y \cos \theta$.

First suppose that there is no applied magnetic field, so that $Y_T = Y_L = 0$ and (6.1.1) becomes

$$n^2 = 1 - X. \hspace{1cm} (6.1.2)$$

The curve relating $n^2$ to $X$ is then a straight line as shown in fig. 6.1 (a). The positive part of the line is re-drawn in fig. 6.1 (b) to represent $\mu^2$, and the negative part, with its sign reversed, is re-drawn in fig. 6.1 (c) to represent $\chi^2$.

* In the first edition of this book there were some incorrect statements about which signs in the equations corresponded to which curves in the diagrams. The statements made there, and here also, about the signs are correct if the refractive index equation is written in the form given in (6.1.1) here. In the first edition, however, it was given in the form of (2.5.14), obtained by dividing the top and bottom of the fraction in (6.1.1) by $(1 - X)$, and the wrong sign was attached to the curves in the range $X > 1$.

I am much indebted to Dr C. S. G. K. Setty for drawing my attention to this error.
When $X > 1$ the attenuation factor $\chi$ has a finite magnitude, although there are no collisions and therefore no mechanism by which energy can be absorbed from the wave. Since, however, $\chi$ is finite only when $\mu = o$ the wave-fields then vary like

$$\exp(-\omega x/c) \exp(iot),$$

and are of the 'evanescent' type encountered in problems of total internal reflection, and in a wave-guide working 'beyond cut-off'. This type of wave does not transmit energy and does not correspond to absorption of energy.

An expression equivalent to this was deduced more simply for this special case by Lorentz (1909), before the general theory was developed. The curves showing $n^2$, $\mu^2$ and $\chi^2$ as functions of $X$ are shown, for the two special cases when $Y = \frac{1}{2}$ and $Y = 2$, in figs. 6.2 and 6.3. For each value of $Y$ there are now two curves, each corresponding to one of the two signs in the equation. The curve corresponding to the upper sign is drawn continuous and that for the lower sign is dashed. Both are straight lines, cutting the $X$-axis at $X = 1 \pm Y$. When $Y > 1$ the dashed line cuts the axis at a negative value of $X$. For positive values of $X$, which are the only ones occurring in practice, the curve corresponds to values of $\mu$ greater than unity.

Fig. 6.4 shows how the lines of figs. 6.2 and 6.3 would change as $Y$ changes. If $Y$ were to increase from zero to infinity the continuous line would swing round from $BAC$ to $DE$, and the dashed line would swing from $BAC$ through $AO$ to $EAD$. 

**Fig. 6.1.** No magnetic field ($Y = o$).

**Fig. 6.2.** Longitudinal propagation from (6.2.1). $Y < 1$ illustrated by $Y = \frac{1}{2}$. The continuous line corresponds to the upper sign.
6.3. The transverse case

When the wave-normal is perpendicular to the magnetic field (transverse case), so that $Y_L = 0$, $Y_T = Y$ (6.1.1) splits into the two equations

$$n_{(u)}^2 = 1 - X,$$

$$n_{(l)}^2 = 1 - X(1 - X)/(1 - X - Y^2)$$

for the upper and lower signs respectively.
according as \( Y > 1 \) or \( < 1 \), as shown in fig. 6.5 for the case when \( Y = \frac{1}{2} \), and in fig. 6.6 for the case when \( Y = 2 \). It is clear that the propagation of the wave represented by these broken curves depends markedly on the steady magnetic field; it is therefore called the 'Extraordinary' wave.

**Fig. 6.6.** Transverse propagation. \( Y > 1 \) illustrated by \( Y = 2 \). The continuous line corresponds to the upper sign.

### 6.4. The general case (\( Y \) less than unity, e.g. \( Y = \frac{1}{2} \))

When \( Y = \frac{1}{2} \) figs. 6.2 and 6.5 represent the curves for the longitudinal and the transverse cases respectively. As the direction of the field is gradually altered there must be a continuous change from the one set of curves to the other, and the question arises 'What is the form of the curves for some intermediate direction of the magnetic field?' At first sight the nature of this continuous change-over is far from obvious. It will be shown that a clear understanding of its nature leads at once to a simple method of visualising the approximate behaviour for any general direction of the field.

Consider first the chief points in the behaviour of the general case as represented by the equation

\[
 n^2 = 1 - \frac{X}{1 - \frac{1}{2} Y^2/(1 - X) \pm \left[ \frac{1}{2} Y^2/(1 - X)^2 + Y^4 \right]^{1/2}}. \tag{6.1.1}
\]

From this equation it may be shown that, for the lower sign, \( n^2 = 0 \) at \( X = 1 \pm Y \), there is an infinity at \( (1 - Y^2)/(1 - Y^2) \) and \( n^2 = 1 \) at \( X = 1 \). For the upper sign, \( n^2 = 0 \) at \( X = 1 \), as in the absence of an imposed magnetic field.

The degeneration of the general case into the transverse case is accomplished by letting \( Y_L \) tend to zero so that the infinity moves to \( X = 1 - Y^2 \), and the zeros being at \( X = 1 \) and \( X = 1 \pm Y \) for

- A cursory examination of (6.1.1) might have suggested that the value of \( X \) which made \( n^2 = 0 \) would involve \( Y_T \) and \( Y_L \) separately; this is not so, and only \( Y = \sqrt{(Y_T^2 + Y_L^2)} \) is involved.
the upper and lower signs respectively, the agreement with fig. 6.5 is clear.

To investigate the degeneration into the longitudinal case of fig. 6.2 \( Y_T \) is allowed to tend to zero, and it is then found that the upper sign gives a zero at \( X = 1 \), whereas the lower sign gives zeros at \( X = 1 \pm Y \) and an infinity at \( X = 1 \). The curves then take the form shown in fig. 6.7(a). If a change-over is made at the point \( X = 1 \), from the continuous curve for the upper sign to the broken curve for the lower sign, and vice versa, the straight lines of the simple Lorentz theory (fig. 6.2) are reproduced.

The curves of figs. 6.5 and 6.7(a) for transverse and longitudinal propagation are plotted together in fig. 6.8 and are labelled \( T \) and \( L \) respectively. It is clear from this figure how the gradual transition from the transverse to the longitudinal case comes about. An intermediate case for \( \theta = 45^\circ \) is sketched in the figure. The Ordinary wave has a zero at \( X = 1 \) and finishes parallel to

\[
n^2 - 1 = -X/(1-Y_L),
\]

i.e. as though it had passed through \( n^2 = 0 \) at \( X = 1 - Y_L \). The Extraordinary wave has zeros at \( X = 1 + Y \) and \( X = 1 - Y \) and an infinity at \( X = (1 - Y^2)/(1 - Y_L^2) \) and finishes parallel to

\[
n^2 - 1 = -X/(1 + Y_L).
\]

For all angles \( \theta \) the curves lie respectively inside the two differently hatched areas, and it is possible to estimate their course with a fair degree of accuracy without any further calculation.

6.5. The general case (\( Y \) greater than unity, e.g. \( Y = 2 \))

The behaviour when \( Y > 1 \) can be studied by applying the principles previously employed when \( Y < 1 \). It will be illustrated for the special case when \( Y = 2 \). The ‘transverse’ curves of fig. 6.6 and the ‘longitudinal’ curves of fig. 6.7(b) are plotted together in fig. 6.9. Equation (6.1.1) shows that in the general case the curve for the upper sign (the (u)-curve) passes through \( n = 0 \) when \( X = 1 \) and has an infinity at \( X = (1 - Y^2)/(1 - Y_L^2) \). When \( |Y_L| > 1 \) this infinity occurs in the (u)-curve and for positive values of \( X \); when \( |Y_L| < 1 \) it occurs in the curve for the lower sign (the (l)-curve) and for negative values of \( X \). The (u)-curves lie in the regions hatched with vertical lines, and the continuous line represents the (u)-curve appropriate to \( \theta = 30^\circ \), i.e. \( |Y_L| = \sqrt{3} \).

The (l)-curve passes through the value \( n = 1 \) when \( X = 1 \) and through \( n = 0 \) when \( X = 1 + Y \), and lies in the region hatched with horizontal lines as shown by the dashed curve drawn for \( \theta = 30^\circ \).
It is especially to be noted that, when \( |Y_L| > 1 \) the branch of the curve labelled ABC is present. This branch goes to infinity at \( X = (1 - Y^2)/(1 - Y_L^2) \), and as \( X \to \infty \) it approaches asymptotically to the line
\[
n^2 - 1 = X/(|Y_L| - 1).
\]

Fig. 6.9. Propagation in a general direction \( \theta \). \( Y > 1 \) illustrated by \( Y = 2 \). The limiting curves for longitudinal (L) and transverse (T) propagation are shown, and the regions within which the curves for other angles must lie are shown by appropriate cross-hatching for the upper (a) and lower (f) signs respectively. Continuous lines correspond to the upper sign. The curves for \( \theta = 45^\circ \) are sketched.

Its form for a series of values of \( |Y_L| \) is sketched in fig. 6.10. It is clear how it approaches the curve DEF as the propagation becomes more nearly 'longitudinal' and \( |Y_L| \) approaches \( Y \). This branch does not occur if \( |Y_L| < 1 \), i.e. if \( \cos \theta < 1/Y \).

Fig. 6.10. To illustrate the form of the upper branch of the curve when \( Y > 1 \). Here \( Y = 2 \) and the form of the upper branch is sketched for different values of \( |Y_L| \). The curves are asymptotic to the lines shown which pass through \( X = 1 - |Y_L| \) and to the lines \( X = (1 - Y^2)/(1 - Y_L^2) \). The branch does not occur for positive \( X \) if \( |Y_L| < 1 \).

6.6. Ordinary and Extraordinary waves

A wave travelling in an ionised medium in the presence of an imposed magnetic field is split into the two characteristic waves of the magneto-ionic theory. If the curves showing \( n^2(X) \) are drawn, as in this chapter, for the two waves, one is found to differ less than the other from the curve appropriate to the absence of an imposed field. This wave is called the Ordinary wave and the other is called the Extraordinary. In the figures the Ordinary wave is represented by the continuous curves, and the Extraordinary by the dashed curves. Comparison with fig. 6.1 shows that the curve for the Ordinary wave is more nearly like that when there is no imposed field.

The names Ordinary and Extraordinary are taken over from the phenomenon of double refraction in optics. It should, however, be realised that the double refraction encountered in the magneto-
MAGNETO-IONIC THEORY

Ionic theory is more complicated than that encountered in optics. For neither characteristic wave in a magneto-ionic medium do the wave-normal and ray directions coincide, whereas in light waves traversing a crystal the Ordinary wave has these directions coincident; it is only in the Extraordinary wave that they are different.

Except when the wave-normal is exactly along the direction of the imposed field the Ordinary wave is represented by the upper sign, provided there are no collisions. The modifications to be introduced when there are collisions will be discussed in §7.3, and it will be shown that under those conditions further consideration needs to be given to the naming of the waves when \( X > 1 \). It will also be shown that, when there are no collisions, there is some confusion about the longitudinal case. This is noticeable in comparing the longitudinal curves of figs. 6.2 and 6.3 with those of fig. 6.7.

CHAPTER 7

THE POLARISATIONS OF THE WAVES

7.1. Introduction

The polarisations of the two characteristic waves which travel with their wave normals in a given direction are given, in terms of the quantity

\[ R = -H_z/H_x = E_z/E_y \]  

(7.1.1)

by (2.6.7)

\[ R = (-i/Y_L)\left\{ \frac{1}{2} Y_L^2/(1 - X - iZ) \mp \frac{1}{2} Y_L^2/(1 - X - iZ)^2 + Y_L^2 \right\}. \]  

(7.1.2)

The sign of the quantity \( Y_L = (\mu_0 e H_0 \cos \theta)/m \) depends on the sign of the charge \((e)\) and on the direction of propagation \((\theta)\), and (7.1.2) should always be used if the sign of \( R \) is to be evaluated. This question of sign becomes confused in what follows through the multiplication of the factor \((1/Y_L)\) into the term under the square root. When that is done (7.1.2) becomes

\[ R = -i\left(\frac{Y_L^2/2Y_L}{(1 - X - iZ)} \mp \left[ (Y_L^2/4Y_L^2)/(1 - X - iZ)^2 + 1 \right] \right). \]  

(7.1.3)

The quantity

\[ Y_L^2/2Y_L = (\mu_0 e H_0/m \omega) \left( \frac{1}{2} \sin^2 \theta / \cos \theta \right) \]  

(7.1.4)

is important. It is convenient to write

\[ \omega_c \equiv (\mu_0 e H_0/m)f(\theta), \]  

(7.1.5)

where

\[ f(\theta) = \frac{1}{2} \sin^2 \theta / \cos \theta, \]  

(7.1.6)

so that

\[ Y_L^2/2Y_L = \omega_c/\omega, \]  

(7.1.7)

\( f(\theta) \) is represented in fig. 7.1. In terms of these quantities (7.1.3) can be rewritten

\[ R = -i\left[ \frac{1}{(\omega/\omega_c)(1 - X - i\omega/\omega_c)} \pm \left[ \left( 1/(\omega/\omega_c)(1 - X - i\omega/\omega_c) \right)^2 + 1 \right]^{1/2} \right]. \]  

(7.1.8)

In general \( R \) is complex, so that the end of the magnetic field vector \( \mathbf{H} \) describes an ellipse, inclined to the \( Ox \)- and \( Oy \)-axes. These axes are in the wave-front, \( Oy \) being in, and \( Ox \)
MAGNETO-IONIC THEORY

67

Equation (7.1.8) shows that the two characteristic waves, corresponding to the two signs, have different polarisations. If the complex quantities \( R_u \) and \( R_t \) represent the values of \( R \) for the upper and lower signs respectively, then the form of (7.1.8) is such that

\[
R_u R_t = 1. \tag{7.1.9}
\]

From this it follows that the two polarisation ellipses are related as shown in fig. 7.2. The one is obtained from the other by reflecting it in the line Ox, which makes 45° with the principal directions, and reversing the direction of rotation.

7.2. The parameter \( \omega_c \) and the polarisation when \( X = 1 \)

The sign of \( \omega_c \), given by (7.1.5), is determined by the sign of the charge \( e \) and by the direction of the imposed magnetic field. Its magnitude is independent of the wave-frequency and varies with the angle \( \theta \) as shown in fig. 7.1. It is zero for longitudinal propagation and infinite for transverse. It plays an important part in the theory and will be much used in what follows. In particular, it plays an important part in determining the polarisation when \( X = 1 \).

Let the magnitude of \( R \) when \( X = 1 \) be represented by \( R(1) \). Equation (7.1.8) shows that

\[
R(1) = \omega_c / \nu \pm i \left[ 1 - (\omega_c / \nu)^2 \right]^{1/2} \quad (\text{\( \omega_c \) positive}). \tag{7.2.1(a)}
\]

Care must be taken with the sign of \( R \) in this expression. The sign of \( \omega_c \) changes with the sign of \( Y_L \), and it has already been seen that the dependence of \( R \) on this sign has become obscured in going from (7.1.2) to (7.1.3). Equation (7.2.1(a)) is correct for positive values of \( \omega_c \), but for negative values it must be written

\[
R(1) = \omega_c / \nu \mp i \left[ 1 - (\omega_c / \nu)^2 \right]^{1/2} \quad (\text{\( \omega_c \) negative}). \tag{7.2.1(b)}
\]

Since

\[
\omega_c = (\mu_0 H_0 e / m) (\frac{1}{2} \sin^2 \theta / \cos \theta),
\]

it (\( \omega_c \)) will be negative for electrons (\( e \) negative) and when there is a positive component of the imposed magnetic field along the

Fig. 7.2. The electric-field polarisations in the plane of the wave front. Ox and Oy are the principal directions and the projection of the imposed magnetic field is along Oy. The positive wave-normal is directed into the paper, along positive Ox. The Ordinary wave ellipse (continuous) and the Extraordinary wave ellipse are related as shown.
positive direction of the wave-normal, as when a wave passes downwards through the ionosphere in the Northern hemisphere. $|R(1)|$ and arg $R(1)$ are plotted in fig. 7.3 as functions of $|\omega_c/v|$ for this case, where $\omega_c$ is negative, and the corresponding ellipses are sketched for the upper and lower signs respectively. If $\omega_c$ were positive, as in the Southern hemisphere, arg $R(1)$ would be altered by $180^\circ$.

![Fig. 7.3. To show the form of the electric-field polarisation ellipse in the plane of the wave-front when $X = 1$ for different values of $|\omega_c/v|$. The coordinates are as in fig. 2.2 and the polarisation ellipse is given in terms of $|R(1)|$ and arg $R(1)$, where $R(1) = E/E_v$ for the upper sign when $X = 1$. It is supposed that there is a positive component of the imposed magnetic field along the positive wave-normal and that the active charges in the magneto-ionic medium are negative. The corresponding ellipses for the upper and lower signs, as seen by an observer looking along the positive wave-normal, are sketched.](image)

Equations (7.2.1) and fig. 7.3 show that, if $|\omega_c/v| < 1$, $R(1)$ is complex, with $|R(1)| = 1$, so that the polarisation ellipses are both contained in a square with its sides parallel to the principal directions. The major axes of the ellipses are inclined at equal but opposite angles to the diagonal of the box, and the senses of rotation are opposite. If $|\omega_c/v| = 1$, $|R(1)| = 1$ and the two ellipses degenerate into a single straight line at $45^\circ$ to the principal directions. If $|\omega_c/v| > 1$,

$$R(1) = \omega_c/v \pm [(\omega_c/v)^2 - 1]^{1/2} \quad (\omega_c \text{ positive}) \quad (7.2.3(a))$$

$$= \omega_c/v \pm [(\omega_c/v)^2 - 1]^{1/2} \quad (\omega_c \text{ negative}) \quad (7.2.3(b))$$

and is real. The two polarisations are represented by straight lines which make equal angles with the bisector of the angle between the principal directions.

### 7.3. Ordinary and Extraordinary waves and the choice of sign in the equations

The magneto-ionic theory, as developed in this book, purports to describe only the nature of the characteristic waves travelling in a homogeneous medium and, for this purpose, it is sufficient to describe the two waves by stating whether the upper or the lower sign is taken in the equations. In practice, however, the theory is often used to discuss the propagation of waves in an inhomogeneous medium, and in particular in one in which the collision frequency and imposed magnetic field are approximately constant but the electron-density changes. This problem can be solved properly only by a full wave analysis, outside the scope of this book. An outline of the results of this analysis is given in § 7.1.9, where it is shown that, if the electron density changes sufficiently slowly then a wave which is a characteristic wave in one place remains a simple characteristic wave as it travels to another place. If, therefore, the refractive index $n$ and the wave-polarisation $R$ were plotted as functions of $X$, the two quantities would be expected to change continuously as $X$ increased, corresponding to the continuous change in the properties of the characteristic wave.

In the discussion of the problem we shall consider the polarisation $R$ as calculated from (2.6.7) (or 7.1.8), and the refractive index $n$ as calculated from

$$n^2 = 1 - \frac{X(1 - X - iz)}{(1-iZ)(1-X-iZ) - \frac{1}{2} Y^2 \pm \left( \frac{1}{2} Y^2 + Y^2(1-X-iZ) \right) Y}$$

(7.3.1)

Equation (7.3.1) corresponds to (6.1.1) and is obtained from (2.6.10) by multiplying top and bottom of the fraction by $(1 - X - iz)$. If now, $R$ and $n$ are calculated for different values of $X$, keeping to one sign in the equation, it is found, under certain circumstances, that the values are discontinuous as $X$ passes through the value unity. An example is shown in Fig. 7.4.

Booker (1934) investigated this phenomenon and showed that the following rules apply:
(a) if \(|\omega/\nu| > 1\) then \(n\) and \(R\) are continuous as \(X\) increases through \(X = 1\) if the same sign in the equation is used throughout, but

(b) if \(|\omega/\nu| < 1\) then \(n\) and \(R\) are continuous as \(X\) increases through \(X = 1\) only if the sign is changed as \(X\) passes through \(X = 1\).

The example in fig. 7.4 falls into the class labelled (b). If Booker's rules are used it is possible to plot continuous curves to show \(n(X)\) and \(R(X)\) for all values of the parameter \(|\omega/\nu|\). Curves of this kind are plotted for \(R(X)\) in later parts of this chapter, but it is first convenient to consider the transition through \(X = 1\) a little more fully.

It was explained in §6.6 that, when there were no collisions \((\nu = 0)\) the two characteristic waves were named Ordinary for the upper sign and Extraordinary for the lower sign. It is convenient to extend this nomenclature to the general case, when \(\nu \neq 0\), as follows. When \(X < 1\) the wave represented by the upper sign is still called the Ordinary. If \(n(X)\) and \(R(X)\) are then plotted as continuous functions of \(X\), over a range which extends from \(X < 1\) to \(X > 1\), the curve which would correspond to the Ordinary wave for \(X < 1\) is said still to correspond to the Ordinary wave for \(X > 1\). Application of Booker's rules then leads to the following results:

The Ordinary wave is represented by the upper sign when \(X < 1\) and by (a) the upper sign when \(X > 1\) if \(|\omega/\nu| > 1\), or (b) the lower sign when \(X > 1\) if \(|\omega/\nu| < 1\).

Application of these rules clarifies the difficulty encountered in §6.6 and illustrated by a comparison of fig. 6.2 with 6.7 (a), and 6.3 with 6.7 (b). Attention was there drawn to the fact that, when propagation was longitudinal, the Ordinary wave was represented in two different ways by the continuous lines in the two figures. The difficulty arises because, for the longitudinal case \(\omega_c = 0\), and for no collisions \(\nu = 0\), so that \(\omega/\nu\) is indeterminate. The correct result can only be deduced by considering limiting cases. If it is supposed that \(\nu = 0\) but that propagation is not quite longitudinal, then \(|\omega/\nu| > 1\), and it is correct to keep to the same sign throughout, as in fig. 6.7. If, however, it is supposed that propagation is exactly longitudinal so that \(\omega_c = 0\), but that \(\nu\) is not quite zero, then \(|\omega/\nu| < 1\) and a single characteristic wave is represented by the curves of fig. 6.7 only if a change is made from one to the other as \(X\) passes through 1. But this gives the same result as figs. 6.2 or 6.3.

7.4. Curves of \(R(X)\)

Scott (1950 a) and Snyder & Helliwell (1952) have published curves which show \(R(X)\) over a wide range of conditions. Their results are here represented differently in fig. 7.5, where \(|R(X)|\) and \(\arg R(X)\) are plotted for the Ordinary wave when the longitudinal component of the magnetic field is along the positive direction of the wave-normal, and the free charges have a negative sign, as for electrons, so that \(\omega_c\) is negative. The abscissa are normalised and represent the quantity \(\xi = (|\omega/\omega_c|) (X - 1)\), and the individual curves correspond to different values of \(|\omega/\nu|\).

Fig. 7.6 shows some ellipses sketched from the data contained in fig. 7.5. The following interesting points are noticeable. The point \(\xi = 0\) corresponds to \(X = 1\), and the corresponding polarisations are those already discussed in §7.2. Negative values of \(\xi\) correspond to \(X < 1\) and \(\xi = -|\omega/\omega_c|\) corresponds to \(X = 0\). The changes of polarisation occur more rapidly with \(X\) when \(|\omega/\omega_c|\) is large, i.e. for high frequencies and for more nearly longitudinal conditions.
When \(|\omega c/\nu| < 1\) it is seen that
\[
|R(\xi)| = (|R(-\xi)|)^{-1}
\]
and
\[
\arg R(\xi) = \arg R(-\xi),
\]
so that the direction of rotation is the same for all values of \(\xi\) and, for \(+\xi\) and \(-\xi\), the 'box' into which the ellipse fits is similar, but rotated through \(\frac{1}{2}\pi\). If \(|\omega_c/\nu|\) is only slightly less than unity the

![Graph](image)

Fig. 7.5. The Ordinary-wave polarisation \(R = -H_y/H_x = E_y/E_x\). The axes are as shown in fig. 2.2 and the effective charges are supposed to be negative. The curves are labelled with the values of \(|\omega_c/\nu|\) for which they apply.

Polarisations of Waves

'box' becomes narrow for values of \(X\) just greater or just less than unity.

For \(|\omega_c/\nu| > 1\) it is seen that
\[
|R(\xi)| = |R(-\xi)|
\]
and
\[
\arg R(\xi) = -\arg R(-\xi),
\]
so that the direction of rotation is opposite for positive and negative values of \(\xi\), but the 'box' into which the ellipse fits is the same for \(+\xi\) and \(-\xi\).

It is important to notice that the major axes of all the ellipses lie in the same pair of quadrants, which, when \(\omega_c\) is negative, are as

![Graph](image)

Fig. 7.6. Sketches of the electric-field polarisation ellipses corresponding to some given values of \(|\omega_c/\nu|\) and \(\xi\). The positive wave-normal direction and the longitudinal component of the imposed magnetic field are directed into the paper. The projection of the imposed magnetic field is shown labelled \(H_T\). With negatively charged electrons the continuous line represents the Ordinary wave and the dashed line the Extraordinary.
shown in fig. 7. This implies that characteristic waves emerging from the ionosphere in the Northern hemisphere would all have their major axes in the N.E.-S.W. pair of quadrants. Measurements have usually, but not always, confirmed this expectation. The exceptions have been explained in terms of ‘coupling’ between the two characteristic waves (see § 17.5).

7.5. Limiting polarisation

From (7.1.8) a finite value of $R$ may be calculated to correspond to a given value of $v$ when $X = 0$, i.e. when the electron density is zero. It might be thought that this would represent the polarisation of a characteristic wave emerging from the ionosphere into uncharged air and that the magnitude of $\xi$ at the level where $X$ was first equal to zero could be determined by measuring this polarisation. Full wave theory, sketched in § 17.6, shows that the matter is not quite so simple. It is, nevertheless, interesting to consider the magnitude of $R$ deduced from (7.1.7) for different values of $v$ when $X = 0$. For this purpose magnitudes of $R(o)$ and $\arg R(o)$ can be read from fig. 7.5 by putting $\xi = -|\omega/\omega_0|$.

CHAPTER 8

THE QUASI-LONGITUDINAL ($QL$) AND QUASI-TRANSVERSE ($QT$) APPROXIMATIONS

8.1. The approximations

The complete expressions

$$n^2 = 1 - \frac{X}{1-iZ - \frac{1}{2}Y_Z^2/(1-X-iZ) \pm \left[ \frac{1}{2}Y_Z^2/(1-X-iZ)^2 + Y_L^2 \right]^{1/2}}$$

and

$$R = \frac{-i}{Y_L} \frac{1}{1-X-iZ} \pm \left( \frac{1}{(1-X-iZ)^2 + Y_L^2} \right)$$

for the complex refractive index and the polarisation are quite complicated. In applying them to problems of propagation through the ionosphere it is convenient to make use of two approximations, one of which applies when waves are propagated sufficiently nearly along the direction of the imposed field and the other when they are propagated sufficiently nearly perpendicular to it. They are called the quasi-longitudinal ($QL$) and the quasi-transverse ($QT$) approximations, and the conditions for which they hold are

$$Y_Z^2/4Y_L^2 \ll |(1-X-iZ)^2| \ldots [QL]$$

and

$$Y_Z^2/4Y_L^2 \gg |(1-X-iZ)^2| \ldots [QT].$$

This chapter will be devoted to a discussion of the circumstances under which one or other of these approximations is valid.

Under the two conditions (8.1.1) and (8.1.2) assume the following approximate forms, the polarisation $R$ being that which corresponds to negative values of $Y_L$ as for a wave passing downwards through the ionosphere in the Northern hemisphere.

$QL$:

$$n_{QL}^2 = 1 - X/(1-iZ \pm |Y_L|),$$

$$R_{QL} = \mp i.$$  

$QT$ (upper sign):

$$n_{QT}^2 = 1 - X/(1-iZ + (1-X-iZ)\cot^2 \theta),$$

$$R_{QT} = 0.$$  

(8.1.7)  

(8.1.8)
MAGNETO-IONIC THEORY

\( QT \) (lower sign):
\[
Q_{T0}^2 = 1 - X[(1 - iZ - Y^2/(1 - X - iZ)], \quad (8.1.9)
\]
\[
R_{T0}^2 = \infty. \quad (8.1.10)
\]

The \( QT \) expression when the upper sign is used is derived as follows. From equation (8.1.1)
\[
n_0^2 = 1 - Y[(1 - iZ - Y^2/(1 - X - iZ)]
\]
\[
+ \left\{ Y^2/(1 - X - iZ) + Y^2 \right\}; \quad (8.1.11)
\]
\[
1 - X[(1 - iZ - Y^2/(1 - X - iZ)]
\]
\[
\times (1 - [1 + Y^2(1 - X - iZ) / Y^2]), \quad (8.1.12)
\]
which, with the conditions of the \( QT \) approximation, gives
\[
n_0^2 = 1 - X[(1 - iZ - Y^2/(1 - X - iZ)](1 - 2 Y^2(1 - X - iZ) / Y^2)]
\]
\[
= 1 - X[(1 - iZ + (1 - X - iZ)(Y^2 / Y^2)]
\]
\[
= 1 - X[(1 - iZ + (1 - X - iZ) \cot^2 \theta], \quad (8.1.13)
\]
where \( \theta \) is the angle between the wave-normal and the imposed magnetic field. A mistake is often made in quoting this result, and the term \( (1 - X - iZ) \cot^2 \theta \) is omitted. The correct form was given by Whitehead (1952).

The expressions (8.1.6), (8.1.8) and (8.1.10) for \( R \) show that when the \( QL \) approximation is valid the two characteristic waves are circularly polarised and when the \( QT \) approximation is valid they are linearly polarised along the principal directions.

8.2. Conditions which satisfy the approximations when \( X = 1 \)

If \( \theta \) is the angle between the wave-normal and the earth’s magnetic field the conditions
\[
Y^2 / 4 Y^2 \leq |(1 - X - iZ)^2| \quad (8.1.3)
\]
for the two approximations can be written
\[
\omega^2 / \omega^2 \leq (1 - X)^2 + Z^2, \quad (8.2.1)
\]
where \( \omega_c \) has already been defined by (7.1.5).

It will be seen that the conditions depend, not only on the gyrofrequency \( (\omega_c) \) and the angle \( (\theta) \) of propagation, but also on the magnitudes of \( X \) and \( Z \). The question of whether a given wave can be discussed in terms of the \( QL \) approximation is not, therefore, simply a question of how nearly the direction of propagation coincides with the direction of the applied field, but depends also on the electron density and the collision frequency; indeed, circumstances can arise in which the \( QL \) approximation holds for wave-normal directions which depart considerably from the direction of the field.

If \( Z \) is taken as constant, and \( X \) is allowed to vary, the condition (8.2.1) for the \( QL \) approximation is most likely to break down when \( X = 1 \), and if it is to hold even then we must have
\[
\omega_c^2 / \omega^2 \leq Z^2
\]
\[
\leq \nu^2 / \omega^2, \quad (8.2.2)
\]
or
\[
\omega_c^2 \leq \nu^2.
\]
If this condition applies at \( X = 1 \) then the \( QL \) approximation can be used there and, a fortiori, for all other values of \( X \). Since \( \omega_c \) depends only on the strength of the imposed field and the direction of propagation the condition is very simple.

When numerical cases are considered it is a reasonably good approximation to assume that the inequalities \( \geq \) and \( \leq \) are satisfied if the larger quantity is nine times the smaller, so that the two approximations require
\[
\nu > 3 |\omega_c| \quad \text{or} \quad Z > 3 Y |f(\theta)| \quad \text{for} \quad QL \quad \text{at} \quad X = 1 \quad (8.2.3)
\]
and
\[
\nu < \frac{1}{3} |\omega_c| \quad \text{or} \quad Z < \frac{1}{3} Y |f(\theta)| \quad \text{for} \quad QT \quad \text{at} \quad X = 1. \quad (8.2.4)
\]

8.3. The two approximations when \( X \neq 1 \)

(a) The \( QT \) approximation

If \( \nu < \frac{1}{3} |\omega_c| \) the \( QT \) approximation is appropriate when \( X = 1 \); it is now necessary to investigate the range of \( X \), near \( X = 1 \), for which it also holds. When \( \nu < \frac{1}{3} |\omega_c| \), the inequality (8.2.1) for the \( QT \) approximation can be written
\[
\omega_c^2 / \omega^2 \geq (1 - X)^2, \quad (8.3.1)
\]
which is equivalent to
\[
Y |f(\theta)| > 3 |(1 - X)|. \quad (8.3.2)
\]
This expression then determines the range of values of $X$, near $X = 1$, for which the QT approximation is appropriate when $\nu < \frac{1}{2} |\omega_e|$. 

(b) The QL approximation

Even if $\nu < \frac{1}{2} |\omega_e|$, so that the QT approximation holds near $X = 1$, the QL approximation may hold for other values of $X$. It will hold, whatever the value of $\nu$, provided

$$Y |f(\theta)| < \frac{1}{2} |(1 - X)| \quad (8.3.3)$$

It is important to find the conditions which make the QL approximation hold when $X = 1 \pm Y_L$. The inequality (8.3.3) then gives, with $\frac{1}{2} \sin^2 \theta / \cos \theta$ written for $f(\theta)$

$$\frac{1}{2} Y \sin^2 \theta / \cos \theta < \frac{1}{2} Y \cos \theta \quad (8.3.4)$$

or

$$\tan^2 \theta < \frac{1}{2}$$

or

$$\theta < 40^\circ. \quad (8.3.5)$$

Provided $\theta < 40^\circ$ it is thus always safe to assume that the QL approximation is valid when $X = 1 \pm Y_L$. If the QT approximation were to be valid when $X = 1 \pm Y_L$, (8.3.2) would require $\tan^2 \theta > 6$ or $\theta > 68^\circ$.

8.4. The QL and QT regions in a model ionosphere

In the terrestrial ionosphere both $X$ and $Z$ vary with height, and it is important to discuss where the QL and QT approximations may separately be used. For the purpose of this discussion use will be made of a model ionosphere described in detail in §13.1. The electron density in this model is represented, in fig. 8.1, in terms of its angular plasma frequency $\omega_p \equiv (4\pi Ne^2/\epsilon_0 m)^{1/2}$ by the line $CD$. This line represents the height at which $X = 1$ for waves of angular frequency $\omega_e$, or actual frequency $f$, as shown by the two separate scales. The collision frequency $\nu$ is represented as a function of height by the line $AB$.

The quantity $\omega_e$ is given the value $4.7 \times 10^5$ sec.$^{-1}$ appropriate to a place where $\theta = 20^\circ$ and $f_H = 1.25$ Mc./s. The horizontal lines $HK$ and $LM$ mark the levels where $\nu = \frac{1}{4} \omega_e$ and $3 \omega_e$, respectively. Then along $CE$, $X = 1$ and $\nu < \frac{1}{4} \omega_e$, so that the QT approximation holds, and along $FD$, $X = 1$ and $\nu > 3 \omega_e$ and the QL approximation holds. Away from the line $X = 1$ the limits within which the two approximations hold are given by

$$\left(\frac{\omega_e}{\nu}\right)^2 \left| \begin{array}{c} < \frac{1}{2} \\ > 9 \end{array} \right| (1 - X)^2 + \frac{1}{2}$$

or

$$\left(\frac{\omega_e}{\nu}\right)^2 \left| \begin{array}{c} < \frac{1}{2} \\ > 9 \end{array} \right| \left(\omega_e - \frac{\nu}{\nu}\right)^2 + \frac{1}{2} \right.$$
8.5. The relation between the QL expression and that for the case of no applied field

When there is no applied magnetic field the general magneto-ionic equation takes the simple form

\[ n^2 = 1 - X/(1 - iZ), \]  

(8.5.1)

and when the QL approximation is valid it takes the form (8.1.5)

\[ n_{QL}^2 = 1 - X/(1 \pm |Y_L| - iZ). \]  

(8.5.2)

Now if we write

\[ X'_{QL} = X/(1 \pm |Y_L|), \]

\[ Z'_{QL} = Z/(1 \pm |Y_L|), \]

(8.5.2) can be put in the form

\[ n_{QL}^2 = 1 - X'_{QL}/(1 - iZ'_{QL}), \]  

(8.5.3)

which is the same as (8.5.1). This transformation will prove convenient in the next chapter.

---

CHAPTER 9

THE EFFECT OF COLLISIONS ON THE CURVES FOR \( \mu \) AND \( \chi \)

9.1. Introduction

When electrons collide with heavy particles the values of the refractive index (\( \mu \)) are modified, and the wave suffers absorption of magnitude represented by \( \chi \). In this chapter it will be shown how the curves for \( \mu(X) \) and \( \chi(X) \) are modified when collisions occur. The effect of collisions on the polarisation of the waves has already been discussed in Chapter 7.

The calculations will first be made on the assumption that there is no applied field. It was shown in §8.5 that, when the QL approximation was valid, the expressions for \( \mu \) and \( \chi \) in the presence of an imposed magnetic field could be transformed into (8.5.1), which gives \( \mu \) and \( \chi \) in the absence of a field, if \( X \) was replaced by \( X'_{QL} = X/(1 \pm |Y_L|) \) and \( Z \) by \( Z'_{QL} = Z/(1 \pm |Y_L|) \). The conclusions for no applied field will therefore be applicable to the QL form of the expressions if these substitutions are made. When \( |Y_L| > 1 \) the values of \( X'_{QL} \) and \( Z'_{QL} \) are negative. It will be found that \( Z \) appears in the expressions only as \( Z^2 \), so that negative values have no special significance; it will, however, be necessary to consider both positive and negative values of \( X \). No simple treatment seems to be possible for the general curves of the magneto-ionic theory, or for those of the QT approximation.

9.2. Curves for \( \mu^2(X) \) and \( \chi^2(X) \) (general)

In the absence of an applied field we have (8.5.1)

\[ (\mu - i\chi)^2 = 1 - X/(1 - iZ). \]  

(9.2.1)

Let us write

\[ (\mu - i\chi)^2 = M - iN, \]  

(9.2.2)

so that

\[ \mu^2 = \frac{1}{2}(M^2 + N^2) + M, \]  

(9.2.3)

\[ \chi^2 = \frac{1}{2}(M^2 + N^2) - M. \]  

(9.2.4)
MAGNETO-IONIC THEORY

The quantities $M$ and $N$ are given from (9.2.1) and (9.2.2)

\[ M = 1 - X/(1 + Z^2) = 1 - \zeta X, \quad (9.2.5) \]
\[ N = XZ/(1 + Z^2) = \zeta XZ, \quad (9.2.6) \]
where

\[ \zeta = 1/(1 + Z^2). \quad (9.2.7) \]

Equations (9.2.3) and (9.2.4) lead immediately to

\[ \mu^2(X) = \frac{1}{2} [(1 + \zeta X(X - 2))^{\frac{1}{2}} + (1 - \zeta X)], \quad (9.2.8) \]
\[ \chi^2(X) = \frac{1}{2} [(1 + \zeta X(X - 2))^{\frac{1}{2}} - (1 - \zeta X)], \quad (9.2.9) \]

from which it is clear that $Z$ appears only through the combination $\zeta = 1/(1 + Z^2)$, and moreover that $\mu(X)$ and $\chi(X)$ are not dependent on the sign of $Z$. As $Z$ takes all values from 0 to $+\infty$, $\zeta$ changes from 1 to 0, and the relation between the two quantities is shown in fig. 9.1.

Figs. 9.2 and 9.3 show curves of $\mu^2(X)$ and 9.4 and 9.5 curves of $\chi^2(X)$ calculated from (9.2.8) and (9.2.9) for a series of different values of $\zeta$. In the plotting of these curves the following facts which are simply deducible from (9.2.8) and (9.2.9) are useful:

(a) $\mu^2$ has a minimum value at $X = 2$ given by

\[ \mu^2(2) = 1 - \zeta. \quad (9.2.10) \]

(b) $\chi^2(2) = \zeta. \quad (9.2.11) \]

(c) $\mu^2(1) = \frac{1}{2} [(1 - \zeta)^{\frac{1}{2}} + (1 - \zeta)]. \quad (9.2.12) \]

(d) $\chi^2(1) = \frac{1}{2} [(1 - \zeta)^{\frac{1}{2}} - (1 - \zeta)]. \quad (9.2.13) \]

In figs. 9.2-9.5, the broken lines with arrows on them indicate how the different parts of the curves move when $Z$ increases steadily ($\zeta$ decreases). It will be noticed that, over some portions of the curves, $\mu^2(X)$ or $\chi^2(X)$ increase or decrease monotonically as $\zeta$ decreases, but over other parts of the curves $\mu^2$ or $\chi^2$ pass through a maximum as $\zeta$ decreases. This phenomenon deserves investigation, as follows.

Differentiation of (9.2.8) and (9.2.9) shows that for a constant value of $X$ a maximum is reached for a value of $\zeta$ given by $\zeta_m$, where

\[ \zeta_m = (X - 4)/(4(X - 2)). \quad (9.2.14) \]
This relation is shown in the curve of fig. 9.6. In the derivation of (9.2.14) the difference of sign between the expression for \( I^2 \) and \( X^2 \) has been lost in a process of squaring, so that it is not clear whether the maximum is in \( I \) or \( X \); it is, however, a simple matter to decide which part of the curve refers to which quantity, and the curve relating \( \zeta_m \) to \( X \) may be labelled as shown in fig. 9.6. Since \( \zeta \) is restricted to the range \( 0 < \zeta < 1 \) the relevant parts of the curve are those shown in heavy lines.

Substitution of the expression for \( \zeta_m \) in the range where \( X < \frac{3}{2} \), shows that the maximum value \( \chi_m(X) \) which \( \chi^2(X) \) ever reaches, for any value of \( \zeta \), is given by

\[
\chi_m^2(X) = \frac{X^2}{8(2-X)}. \tag{9.2.15}
\]

This is plotted as the curve \( ABC \) in fig. 9.7. In the same way, when \( X > 4 \), it can be shown that the maximum value \( \mu_m(X) \) which \( \mu^2(X) \) reaches, for any value of \( \zeta \), is given by

\[
\mu_m^2(X) = \frac{X^2}{8(X-2)}, \tag{9.2.16}
\]

and this is plotted as the curve \( DE \) in fig. 9.7. The shaded regions are those in which \( \mu^2(X) \) and \( \chi^2(X) \) respectively pass through a maximum if \( X \) is kept constant and \( \zeta \) is varied smoothly from 1.

---

**Fig. 9.4.** Curves of \( \chi^2(X) \) for different values of \( \zeta \) (or \( Z \)). The part near the origin is shown in more detail in fig. 9.5.

**Fig. 9.5.** Curves of \( \chi^2(X) \) for different values of \( \zeta \) (or \( Z \)). Curves on a larger scale are shown in fig. 9.4.

**Fig. 9.6.** The values (\( \zeta_m \)) of \( \zeta \) which give maximum values of \( \mu \) or \( \chi \).

Here \( \zeta_m = (X-4)/4(X-2) \).

**Fig. 9.7.** The maximum values (\( \mu_m, \chi_m^2 \)) of \( \mu^2 \) and \( \chi^2 \) which are reached for the appropriate values of \( \zeta_m \) given in fig. 9.6.
to 0. The curves for $\mu^2(X)$ and $\chi^2(X)$ respectively never rise above $DE$ and $ABC$. Inspection of figs. 9.4 and 9.5 jointly shows that as $\zeta$ decreases from 1 to $\frac{1}{2}$ the $\chi^2(X)$ curve touches the curve $ABC$ at a point which moves from $A$ past $B$ and $C$ to $-\infty$. As $\zeta$ continues to decrease from $\frac{1}{2}$ to 0 the $\mu^2(X)$ curve touches $DE$ at a point which moves in $\infty$ to $D$.

The way in which the curves change with $\zeta$ is shown in figs. 9.2-9.5 by the arrows on the broken lines which indicate how the curves move as $Z$ increases steadily from 0 to $\infty$ so that $\zeta$ decreases steadily from 1 to 0. In some regions of the diagrams the corresponding values of $\mu^2$ (or $\chi^2$) pass through a maximum, in agreement with what was said above. Figs. 9.2 and 9.4 show the variations over a wide range of $X$, and figs. 9.3 and 9.5 show a more restricted range in more detail. In the interpretation of these curves the relation between $\zeta$ and $Z$ shown in fig. 9.1 should be borne in mind.

Although the figures relate to the case with no applied field, they can, in addition, be applied, as explained in §9.1, to the QL approximation as follows:

(a) If the curves for negative values of $X$ are reflected in the line $X = 0$ they correspond, with suitable change of scale, to the negative sign in the expression (8.5.2) for the longitudinal case when $|Y_L| > 1$.

(b) The curves for positive values of $X$ correspond to the positive sign in the expression for the longitudinal case; they also correspond, with a different change of scale, to the negative sign in the expression for the longitudinal case, when $|Y_L| < 1$.

9.3. The detailed changes of $\mu$ and $\chi$

In this section attention is drawn to some of the details shown in the curves for $\mu^2(X)$ and $\chi^2(X)$.

(a) The curves for $\mu^2(X)$. Figs. 9.2 and 9.3.

(c) For $\zeta = 1$ ($Z = 0$) the curve for $\mu^2(X)$ is $ABC$ and for $\zeta = 0$ ($Z = \infty$) it is $DE$ (i.e. $\mu^2(X) = 1$).

(d) There is a minimum of $\mu^2$ at $X = \pm 2$, and its value is $1 - \zeta$.

(e) For $X < 4$ a decrease of $\zeta$ (increase of $Z$) causes the value of $\mu$ to increase (for $X > 1$) or decrease (for $X < 1$) so that the curve falls more nearly along $\mu^2 = 1$.

(f) For $X > 4$ a decrease of $\zeta$ (increase of $Z$) produces in $\mu$ first an increase from zero to a value greater than 1 and then a decrease to the value 1 for small $\zeta$ (large $Z$).

(b) The curves for $\chi^2(X)$. Figs. 9.4 and 9.5.

(g) For $\zeta = 1$ ($Z = 0$) the curve for $\chi^2(X)$ is $GHI$ and for $\zeta = 0$ ($Z = \infty$) $\chi^2(X)$ is zero.

(h) For negative values of $X$ (extraordinary longitudinal wave with $|Y_L| > 1$) $\chi = 0$ for both extreme values of $\zeta$.

(i) For $X > \frac{3}{2}$ a decrease of $\zeta$ (increase of $Z$) decreases the value of $\chi^2(X)$ steadily.

(j) For $X < \frac{3}{2}$ (including all negative values) a decrease of $\zeta$ (increase of $Z$) produces in $\chi$ first an increase from zero to a maximum, followed by a decrease to zero.

9.4. Approximations for small values of $Z$, and the conditions for reflection

In cases of practical importance it often happens that $Z^2 \ll 1$, and it is then of interest to investigate the forms which the curves take.

If $Z^2 \ll 1$ (say $Z < \frac{1}{2}$) $\zeta = 1 - Z^2$ and (9.2.8) and (9.2.9) lead to the following values of $\mu^2$ and $\chi^2$ except near $X = 1$:

$X > 1$ $\mu^2 = \frac{1}{2}Z^2X^2/(X-1)$, \hspace{1cm} (9.4.1)

$X < 1$ $\mu^2 = (1-X)+\frac{1}{2}XZ^2(4-3X)/(1-X)$, \hspace{1cm} (9.4.2)

$X > 1$ $\chi^2 = (X-1)+\frac{1}{2}XZ^2(4-3X)/(X-1)$, \hspace{1cm} (9.4.3)

$X < 1$ $\chi^2 = \frac{1}{2}Z^2X^2/(1-X)$. \hspace{1cm} (9.4.4)

These equations show that $\mu^2$ and $\chi^2$ can be expressed as follows, except near $X = 1$:

$\mu^2(X, Z) = \mu^2(X, 0)+Z^2F(X)$, \hspace{1cm} (9.4.5)

$\chi^2(X, Z) = \chi^2(X, 0)+Z^2G(X)$, \hspace{1cm} (9.4.6)

Here $F(X)$ and $G(X)$ are the functions shown in fig. 9.8. It can be seen that $F(X)$ has a minimum at $X = 2$ so that $\mu^2$ has a minimum there, in conformity with the general rule derived above. It is also seen that $G(X)$ changes sign at $X = \frac{3}{2}$ in conformity with the requirement that, for small values of $Z$, the curve for $\chi^2(X)$ must lie...
above the value \( \chi^*(0) \) for \( X < \frac{1}{2} \) but must lie below for \( X > \frac{1}{2} \) (cf. fig. 9.5).

Equations (9.2.10) and (9.2.11) can be used to show that, when

\[
Z^2 < 1 \quad \begin{cases} \mu^2(Z, Z) = Z^2, \\ \chi^2(Z, Z) = 1 - Z^2. \end{cases}
\]

(9.4.7)
(9.4.8)

Also, from (9.2.12) and (9.2.13)

\[
\mu^2(1, Z) = \frac{1}{2} \left( (1 - \zeta)^\frac{1}{2} \pm (1 - \zeta) \right),
\]

(9.4.9)

so that \( Z^2 < 1 \)

\[
\mu^2(1, Z) = \chi^2(1, Z) = \frac{1}{2} Z.
\]

(9.4.10)

The magnitudes of \( \mu \) and \( \chi \) when \( X = 1 \), given by (9.4.10), are especially important because it frequently happens that a possible reflection level for vertically incident waves is situated near the place where \( X = 1 \). If, at that level, both \( \mu \) and \( \chi \) are sufficiently small, then it is shown in Chapter 17 that appreciable reflection might occur. If, for example, \( Z = 0.1 \), then

\[
\mu(1, Z) = \chi(1, Z) = 0.2,
\]

(9.4.11)

and this condition will be used in what follows to determine whether or not appreciable reflection is to be expected near the level where \( X = 1 \).

It will be noticed that, when \( Z^2 < 1 \), the magnitudes of \( \mu^2 \) and \( \chi^2 \) at \( X = 1 \) are proportional to \( Z \), but that, for \( X \) not near 1, the departures from the magnitudes without collisions are proportional to \( Z^2 \).

9.5. Conditions for reflection when \( Z^2 \gg 1 \) at \( X = 1 \)

If there were no collisions and no imposed magnetic field, a wave travelling vertically into the ionosphere would be reflected at a level where \( X = 1 \). In this section we shall consider what would happen if there were no magnetic field but the collisions were so numerous that \( Z^2 \gg 1 \) at the level where \( X = 1 \). Equations (9.2.8) and (9.2.9) show that, if \( Z^2 \gg 1 \),

\[
\mu^2 = \frac{1}{2} \left( 1 + \frac{X(X-2)}{Z^2} \right)^\frac{1}{2} \pm \frac{1 - X}{Z^2},
\]

(9.5.1)

Near \( X = 1 \) the inequalities \( Z^2 \gg X \) or \( X^2 \) will hold and (9.5.1) gives

\[
\mu^2 = 1 + \left( \frac{1}{2} X^2 - X \right) / Z^2,
\]

(9.5.2)

\[
\chi^2 = \frac{1}{2} X^2 / Z^2
\]

(9.5.3)

so that \( \mu \) is near unity, \( \chi \) is small, there is no reflection, and the wave passes on to higher levels.
As the height increases the electron density (proportional to \(X\)) increases and the collision frequency (proportional to \(Z\)) decreases, so that the ratio \(X/Z\) increases and ultimately a height is reached where \(X = Z\). Let us suppose that this level is above the place where \(X = 3\) so that we may write \(Z^2 = X^3 \gg 1\) and (9.5.1) still applies and gives

\[
\begin{align*}
\mu^2/\chi^2 & = 1/\{1 + (1 + 1)^{4\pm 1}\}, \\
\text{or}^* \quad & \mu(X/Z = 1) = 1, \quad (9.5.4) \\
& \chi(X/Z = 1) = 0.45, \quad (9.5.5)
\end{align*}
\]

so that \(\chi\) has now become appreciable, although \(\mu\) is still not much different from unity. In order to estimate the distance in which \(\chi\) has changed from its small value of (9.5.3) to the appreciable value of (9.5.5) we insert the value \(X/Z = 0.3\) in (9.5.1) to obtain

\[
\begin{align*}
\mu(X/Z = 0.3) & = 1, \quad (9.5.6) \\
\chi(X/Z = 0.3) & = 0.16. \quad (9.5.7)
\end{align*}
\]

It thus appears that, as \(X/Z\) increases from 0.3 to 1, \(\mu\) does not alter much, but \(\chi\) changes from about 0.16 to about 0.45.

Now it is shown in § 17.2 that if \(\mu\) or \(\chi\) changes by an appreciable fraction of itself in a distance comparable with one wave-length, appreciable reflection will take place. It is therefore concluded that, under the conditions here envisaged, appreciable reflection will occur near the level where \(X = Z\).

9.6. Results for the QL approximation

It is of interest to examine the form of the \(\mu^2(X)\) and \(\chi^2(X)\) curves for a series of different values of \(\nu\) when the QL approximation is valid. First consider the case where \(|Y_L| < 1\) so that \(X_{QL}^2\) is positive for both the signs in the equation. The curves for \(\mu^2(X)\) and \(\chi^2(X)\) are then as shown for positive \(X\) in figs. 9.2 and 9.4. The scale of \(X\) has to be altered so that the abscissae marked \(X = 1\) should be re-labelled \(X = (1 + |Y_L|)\) for the two signs respectively.

Consider now what happens as \(\nu\) is increased from zero to a large value. The quantities \(\xi_{QL}^2\), given by

\[
\xi_{QL}^2 = 1/\{1 + v^2/|Y_L|^2\},
\]

* It should be noted that (9.5.4) and (9.5.5) refer to \(\mu\) and \(\chi\) and not to \(\mu^2\) and \(\chi^2\).

decrease from 1 to 0, but at different rates for the two signs. When \(\nu = 0\) (\(\xi' = 1\)) \(\mu\) is represented by the curve \(ABC\); as \(\nu\) increases and \(\xi'\) decreases the curve for \(\mu\) alters as shown by the arrows in fig. 9.2. For values of \(X < \frac{1}{2}(|Y_L|)\) the value of \(\mu\) increases steadily, but for values of \(X > \frac{1}{2}(|Y_L|)\) the value of \(\mu\) first increases, and then finally decreases. When \(\xi'\) is very small (\(\nu\) very large) the value of \(\mu\) is unity for all values of \(X\).

While these changes of \(\mu\) are taking place \(\chi^2(X)\) alters as shown in figs. 9.4 and 9.5. When \(\nu = 0\), \(\chi^2\) is represented by the line \(OHI\), and as \(\nu\) increases (\(\xi'_{QL}\) decreases) the curves for \(\chi\) assume, in succession, the forms shown. For \(X < \frac{1}{2}(|Y_L|)\) the value of \(\chi\) for a given value of \(X\) rises to a maximum for the value of \(\xi'\) given by (9.2.16) and then decreases again to zero when \(\xi' \to 0\) \((\nu \to \infty)\). For \(X > \frac{1}{2}(|Y_L|)\) the value of \(\chi\) decreases steadily as \(\nu\) increases. When \(\nu\) is very large \(\chi = 0\) for all values of \(X\).

Next consider the case when \(|Y_L| > 1\). The behaviour for the upper sign is similar to that for \(|Y_L| < 1\), but for the lower sign \(X_{QL}^2 = X/(1 - |Y_L|)\) is now negative, so that the negative values of \(X\) must be used in the curves of figs. 9.2 and 9.4. From fig. 9.2 it is seen that the value of \(\mu\) is greater than 1 for all values of \(X\). For any given value of \(X\) an increase of \(\nu\) results in a decrease of \(\mu\), and when \(\nu\) becomes very great \(\mu = 1\) for all values of \(X\).

The corresponding changes of \(\chi^2(X)\) are clear from fig. 9.4. When \(\nu = 0\) or \(\infty\), \(\chi^2\) is zero for all values of \(X\). As \(\nu\) increases from zero, \(\chi^2\) for any value of \(X\) first increases and then decreases.

In Chapter 10 it will be of interest to consider the magnitudes \(\mu(1 + Y_L)\) and \(\chi(1 + Y_L)\) which \(\mu\) and \(\chi\) assume when \(X = 1 + Y_L\) and the QL approximation holds. Equation (9.4.11) then shows, if \(X' = 1/((1 + Y_L))\) and \(Z' = Z/((1 + Y_L))\) are substituted for \(X\) and \(Z\) respectively, that, if

\[
Z < 0.1\{(1 + Y_L)\},
\]

then

\[
\mu(1 + Y_L) = \chi(1 + Y_L) < 0.2.\]

Comparison with the results of § 9.2, (9.2.10) shows that the minimum values of \(\mu\) will occur when \(X = 2(1 + Y_L)\) and will have values given by

\[
\mu^2_{\text{min.}} = 1 - \{1 + Z^2((1 + Y_L)^2)\}^{-1}.\]
The results of §9.5 can also profitably be extended to cover the case where a magnetic field is present and the QL approximation holds. Substitution of $X'$ and $Z'$ for $X$ and $Z$ then leads to the conclusion that, for both characteristic waves, if $Z^2 \gg 1$ when $X' = 1$ there will be no reflection at this level but, if $Z'$ decreases and $X'$ increases as the height increases reflection of both waves will occur near the level where $Z' = X'$, i.e. where $Z = X$.

9.7. The QT approximation near $X = 1$, upper sign

If there were no collisions the refractive index for the Ordinary wave (upper sign) would reach zero when $X = 1$ and the wave would be reflected from that height in the ionosphere. If the collision frequency is sufficiently small the QT approximation will be valid near $X = 1$ and the refractive index can become small. The magnitudes of $\mu$ and $\chi$ under these conditions will be considered below.

If, in general, we write

$$n^2 = 1 - 1/(a - ib),$$

(9.7.1)

then it follows that, if $\mu^2 \gg \chi^2$,

$$\mu^2 = 1 - a/(a^2 + b^2),$$

(9.7.2)

$$\chi = \frac{b}{2a} \left(\frac{1}{\mu} - \mu\right).$$

(9.7.3)

Now from (8.1.13) for the QT approximation, upper sign,

$$a = \left(1 - (1 - X) \cot^2 \theta\right)/X,$$

(9.7.4)

$$b = (Z/X) \cosec^2 \theta,$$

(9.7.5)

and, if it is supposed that $(1 - X) \cot^2 \theta < 0.1$ and $Z \cosec^2 \theta < 0.1$, (9.7.2) and (9.7.3) lead to

$$\mu^2 = (1 - X) \cosec^2 \theta,$$

(9.7.6)

and

$$\chi = \frac{1}{2Z} \cosec^2 \theta \left(\frac{1}{\mu} - \mu\right).$$

(9.7.7)

These two expressions no longer hold when $X = 1$ because then it cannot be assumed that $\mu^2 \gg \chi^2$. When $X = 1$, (8.1.13) for the QT approximation with the upper sign shows that

$$n^2(1)_{QT} = 1 - 1/(1 - iZ \cosec^2 \theta),$$

(9.7.8)

so that, if $Z^2 \cosec^4 \theta \ll 1$, an argument like that leading to (9.4.10) shows that

$$\mu^2(1, Z) = \chi^2(1, Z) = \frac{1}{2} Z \cosec^2 \theta,$$

(9.7.9)

and, in particular, if $Z \cosec^2 \theta < 0.1$, or $Z < 0.1 \sin^2 \theta$, then

$$\mu(1, Z)_{QT} = \chi(1, Z)_{QT} < 0.2.$$  (9.7.10)

It is next convenient to express (9.7.7) in terms of the group refractive index ($\mu'$) given, as shown in Chapter 11, by

$$\mu' = \mu + \omega(d\mu/d\omega),$$

(9.7.11)

so that from (9.7.6), with $X = \omega \mu/\omega^2$,

$$\mu' = (1/\mu) \cosec^2 \theta$$

(9.7.12)

and from (9.7.7)

$$\chi = \frac{1}{2Z} (\mu' - \mu \cosec^2 \theta).$$

(9.7.13)

Equations (9.7.10) and (9.7.13) will be useful in §§10.3 and 10.4.
CHAPTER 10
SOME REPRESENTATIVE CURVES FOR THE
REFRACTIVE INDEX (μ) AND THE
ABSORPTION INDEX (χ)

10.1. Introduction
In Chapter 6 curves were sketched to show how the refractive index μ(X) and the absorption index χ(X) depended on X (proportional to N) when there were no collisions. In this chapter corresponding curves will be discussed for cases where collisions are present.

The calculation of μ(X) and χ(X) from (2.6.10) for any given set of parameters is tedious. Some workers have evolved simplified methods of calculation, either by noticing that (2.6.10) can be rewritten in terms of hyperbolic functions of a complex quantity, or by using mechanical linkages. Others have made calculations with the help of electronic digital computers, and there exist a large number of curves, or tables, which give the results of solving the magneto-ionic equation for a wide range of parameters. These are listed in the Bibliography. In this chapter a representative selection of these results is presented in graphical form in such a way as to show up some general principles and to establish a method for sketching the essential features of the curves. By making use of the curves here presented, and of the principles established, it should be possible to sketch the curves appropriate to any other set of parameters.

First in §§10.2 and 10.3 the results of Chapters 8 and 9 will be used to establish some general features of all μ(X) and χ(X) curves. In §10.4 it will be shown how these general features can be recognised in the particular curves chosen.

10.2. The magnitudes of μ and χ near the possible reflection levels
A wave vertically incident on the ionosphere will be reflected at a level where μ = o, provided χ is not too large at and below that level. Now it was shown in Chapter 6 that, when ν = o, the only values of X which make both μ and χ equal to zero are one or more of the values X = (1 — Y), 1, or (1 + Y), so that reflection would occur at one or more of these levels. When the electrons make collisions with heavy particles the value of μ never falls to zero, but it is shown in Chapter 17 that appreciable reflection may still be expected, even then, near any level at which μ decreases sufficiently rapidly to a small value, provided that χ is not too large. If the collision frequency is small enough this decrease will occur near one or more of the levels where X = (1 — Y), 1, or (1 + Y), and attention will therefore be concentrated on the values of μ and χ for these values of X.

10.3. The method for sketching approximate curves
In any one of the curves to be discussed in this chapter the collision frequency (ν), wave-frequency (ω), magnetic field (through Y), and angle of propagation (θ), are assumed constant, and the electron density (proportional to X) is plotted as abscissa.

When any particular set of parameters has been chosen it is convenient first to use the criteria of Chapter 8 to decide whether either the QL or the QT approximation is valid. In general, different approximations will be appropriate for different ranges of X. Portions of the curve are then sketched in by making use of the results of Chapter 6, appropriate to the case when ν = o, using the longitudinal curves over the region where the QL approximation holds and the transverse curve where the QT approximation holds.

It is next necessary to decide whether the values of μ(X, ν) and χ(X, ν) will differ appreciably from the values μ(X, o) and χ(X, o) they would have if ν = o. For this purpose the results of Chapter 9 are used, and in particular it is investigated whether μ can reach small values near X = (1 ± Y) or 1. Some general rules can be stated as follows.

(a) Where X = 1
(i) The QL approximation is valid when X = 1 if

\[ ω > 3 |ω_e| \]  
(see (8.2.4)),

and then the value of μ cannot be small.
(ii) The QT approximation is valid when \( X = 1 \) if
\[
\nu < \frac{1}{2} |\omega e| \quad \text{(see (8.2.5))},
\]
and then the value of \( \mu \) might be small, provided \( \nu \) were small enough.

(iii) When the QT approximation holds \( \mu(1) \) and \( \chi(1) \) are small
\( (<0.2) \) if
\[
Z < 0.1 \sin^2 \theta \quad \text{(see (9.7.10))}.
\]

(iv) Condition (iii) is automatically satisfied through condition
(ii) if
\[
\frac{1}{2} Y \sin^2 \theta/(|\cos \theta|) < 0.3 \sin^2 \theta,
\]
i.e.
\[
Y < 0.6 |\cos \theta|.
\]
\[(10.3.4)\]

(b) Where \( X = 1 \pm Y_L \)

(i) The QL approximation is valid when \( X = 1 \pm Y_L \) provided
\( \theta < 40^\circ \) (see (8.3.5)).

(ii) When the QL approximation holds \( \mu(1 \pm Y_L) \) and \( \chi(1 \pm Y_L) \)
are small \( (<0.2) \) if
\[
Z < 0.1 [(1 \pm Y_L)] \quad \text{(see (9.6.1))}
\]
which, with \( \theta < 40^\circ \), is approximately the same as
\[
Z < 0.1 [(1 \pm Y)].
\]
\[(10.3.6)\]

(c) \( \mu_{\text{min.}} \) for QL approximation

When the QL approximation holds the minimum value of \( \mu \)
occurs when \( X = 2(1 \pm |Y_L|) \) and
\[
\mu_{\text{min.}}^2 = 1 - \frac{1}{1 + Z^2} (1 \pm |Y_L|)^{-1} \quad \text{(see (9.6.2))}.
\]
\[(10.3.7)\]

10.4. Some particular examples

The types of result obtained in general are illustrated by the
curves of \( \mu(X) \) in figs. 10.1–10.5. In studying these curves it should
be remembered that the quantity plotted is \( \mu \), whereas in most of
the rest of the book it is \( \mu^2 \). The curves are drawn for the values
of \( Y, Z \) and \( \theta \) indicated, and have been chosen from the available
computed curves, to illustrate general points. The original sources
of the computations are indicated on the curves.
### Notes

(a) The curves are of the QL or QT type according as $Z > \omega_0 < |\omega_0/\omega|$.  
(b) For the curves of QT type $\mu(t)$ is not small if $Z > \omega_1 \sin^2 \theta$ (see (10.3.3)),  
(e.g. $\theta = 10^\circ$, $Z = \omega_1$) \{ $\theta = 30^\circ$, $Z = \omega_1$, $Z = \omega_2$. \} 
(c) $\mu$ is small near $X = 3$ (i.e. $1 + Y$) when $Z < \omega_3$ (see (10.3.6)).  
Representative conditions to which these curves would apply are  
$$ f_B = 1.25 \text{ Mc./s.}, \quad f = 625 \text{ kc./s.}, \quad \omega = 4 \times 10^6 \text{ sec}^{-1}. $$

### Fig. 10.2. $\mu(X)$ curves for $Y = 2$ so that $\omega_1(1 + Y) = \omega_3$.  

### Notes

(a) The curves are of the QL or QT type according as $Z > \omega_0 < |\omega_0/\omega|$.  
(b) For the curve of the QT type $\mu(t)$ is small only if $Z < \omega_1 \sin^2 \theta$ (see (10.3.3)).  
(c) $\mu$ is small near $X = 4$ (i.e. $1 + Y$) when $Z < \omega_1(1 + Y)$ (see (10.3.6)).  
(d) For curves of the QL type minimum values of $\mu$ calculated from (10.3.7) are shown as dots. They agree with the computations.  
Representative conditions to which the curves would apply are  
$$ f_B = 1.5 \text{ Mc./s.}, \quad f = 500 \text{ kc./s.}, \quad \omega = 3.14 \times 10^6 \text{ sec}^{-1}. $$

### Fig. 10.3. $\mu(X)$ curves for $Y = 3$ and $\theta = 19^\circ$.  
Here $\omega_1(1 + Y) = \omega_4$, $\omega_1 \sin^2 \theta = \omega_0$, $|\omega_0/\omega| = \omega_1$.  

### Notes

(a) The curves are of the QL or QT type according as $Z > \omega_0 < |\omega_0/\omega|$.  
(b) For the curve of the QT type $\mu(t)$ is small only if $Z < \omega_1 \sin^2 \theta$ (see (10.3.3)).  
(c) $\mu$ is small near $X = 4$ (i.e. $1 + Y$) when $Z < \omega_1(1 + Y)$ (see (10.3.6)).  
(d) For curves of the QL type minimum values of $\mu$ calculated from (10.3.7) are shown as dots. They agree with the computations.  
Representative conditions to which the curves would apply are  
$$ f_B = 1.5 \text{ Mc./s.}, \quad f = 500 \text{ kc./s.}, \quad \omega = 3.14 \times 10^6 \text{ sec}^{-1}. $$
Fig. 10.4. $\mu(X)$ curves for $Y = 10$ and $\theta = 25^\circ$. Here $\omega_1(1 + Y) = 1.1$, $\omega_1 \sin^3 \theta = 0.018$, $|\omega|/\omega_1 = 1.06$.

Notes

(a) The upper set of curves are of the $QT$ type ($Z < |\omega_1/\omega|$) and the lower set of the $QL$ type ($Z > |\omega_1/\omega|$).

(b) For the curves of $QT$ type $\mu(1)$ is not small because $Z > \omega_1 \sin^3 \theta$.

(c) $\mu$ is small near $X = 11$ (i.e. $1 + Y$) provided $Z < 1.1 = \omega_1(1 + Y)$ (see (10.3.6)).

Representative conditions to which these curves would apply are

$$f_H = 1.6 \text{ Mc./s., } f = 150 \text{ kc./s., } \omega = 9.4 \times 10^6 \text{ sec}^{-1}.$$
The preceding curves have been plotted from data obtained from the following sources:

- Fig. 10.1, Budden (1958) and Landmark & Lied (1957b).
- Fig. 10.2, Budden (1958).
- Figs. 10.3, 10.4 and 10.5, Ross & Fluke (1953).

It will be noted that, in figs. 10.3, 10.4 and 10.5 greater values of \( Z \) are taken for the lower frequencies, i.e. when \( Y \) is greater. This is because, in the available tabulations, roughly the same range of values of \( v \) was used for all frequencies, so that \( Z (= v/\omega) \) is greater for the lower frequencies.

In the figures, the values of \( Z \) used correspond to the following values of the collision frequency \( v \):

**Fig. 10.1**

\[
\begin{array}{cccc}
\nu (\text{sec}^{-1}) & 0.01 & 0.1 & 0.5 & 1.0 \\
0.6 \times 10^8 & 1.6 \times 10^8 & 8 \times 10^8 & 1.6 \times 10^7 \\
\end{array}
\]

**Fig. 10.2**

\[
\begin{array}{cccc}
\nu (\text{sec}^{-1}) & 4 \times 10^4 & 4 \times 10^5 & 8 \times 10^5 & 2 \times 10^6 \\
4 \times 10^6 & 8 \times 10^6 & \text{(no tabulation)} & \text{(no tabulation)} \\
\end{array}
\]

**Fig. 10.3**

\[
\begin{array}{cccc}
\nu (\text{sec}^{-1}) & 10^4 & 5 \times 10^5 & 10^6 & 10^7 \\
0.032 & 0.16 & 0.32 & 2 & 16 \\
\end{array}
\]

**Fig. 10.4**

\[
\begin{array}{cccc}
\nu (\text{sec}^{-1}) & 10^5 & 5 \times 10^6 & 10^7 \\
0.11 & 0.21 & 0.42 & 0.85 \\
\end{array}
\]

**Fig. 10.5**

\[
\begin{array}{cccc}
\nu (\text{sec}^{-1}) & 3 \times 10^7 & 5 \times 10^6 & 2 \times 10^7 \\
1.9 & 4.5 & 13 & 32 \\
\end{array}
\]

Figs. 10.3, 10.4 and 10.5 illustrate the important fact that, for long waves \((Y > 1)\), \( \mu \) can become small (near \( X = 1 + Y \)) for only one of the two characteristic waves, and even then only if collisions are sufficiently infrequent. Moreover, the wave for which this small value of \( \mu \) occurs can be either the Ordinary (solid line) or the Extraordinary (dashed line) according as \( v > \omega_c \) or \( < \omega_c \). See, for example, the difference between the curves for \( Z = 0.16 \) and 0.32 in fig. 10.3; or \( Z = 1.9 \) and 4.5 in fig. 10.5.

\section*{CHAPTER II}

**THE GROUP VELOCITY**

\subsection*{11.1. No magnetic field}

In a valuable method of ionospheric exploration radio-frequency pulses are timed as they travel through the ionosphere. It is important to calculate the group velocity with which they travel.

The general expression for the group velocity of a wave

\[
E = \mathcal{E} \exp \{i(\omega t - kx)\}
\]

as derived in textbooks on wave motion, or physical optics, is usually given in the form

\[
V_{gp} = d\omega/dk.
\]

By making the substitutions \( \mu' = c/V_{gp}, k = \omega/c \) (11.1.1) may be written

\[
\mu' = \mu + \omega(d\mu/d\omega),
\]

and this form is particularly useful for the present discussion. \( \mu' \) may be called the group refractive index.

If there is no applied magnetic field, and if collisions of electrons with heavy particles may be omitted, the expression for \( \mu \) is (see (2.2.6))

\[
\mu^2 = 1 - \omega_{ci}^2/\omega^2,
\]

and (11.1.2) shows at once that

\[
\mu \mu' = 1.
\]

This simple result leads to some interesting and useful relations. Its application is, however, restricted to those cases where (11.1.3) is approximately true, and that implies that \( \omega^2 > \omega_{ci}^2 \), a condition which is rarely satisfied in the terrestrial ionosphere.

\subsection*{11.2. An imposed magnetic field}

When the refractive index is given by the full equation (2.6.10) of the magneto-ionic theory the expression for the group refractive index is complicated. It has been evaluated, for the case when collisions are absent, by several authors for special cases. For
detailed results the original papers should be consulted; that by Shinn & Whale (1952) is most complete. Here only a few points of special interest will be mentioned.

The paper by Shinn & Whale contains a series of curves which show \( \mu'(X) \) for a series of different frequencies \( (Y) \) when the wave makes an angle \( \theta = 23^\circ 16' \) with the imposed magnetic field, as for vertical propagation in south-east England. Their curves for the Ordinary wave are shown in fig. 11.1. Because \( \mu' \) reaches an infinite value when \( X = 1 \), it is desirable to plot the curve with the scale shown, for which \( (1/\mu') \) is proportional to the distance from the top of the figure. Curves for values of \( Y \) from 0 to 4 are included in the figure. That for \( Y = 0 \) corresponds to the absence of an imposed field and represents the curve \( \mu' = (1 - X)^{-1} \). When there is an imposed magnetic field this curve is appropriate to an infinite frequency. At lower frequencies the curves corresponding to greater values of \( Y \) are appropriate. It will be seen that for small \( X \) the value of \( \mu' \) is always less than that for \( Y = 0 \), but it approaches unity \( \mu' \) is greater than for \( Y = 0 \). It therefore follows that as an Ordinary wave-group ascends in the ionosphere to the level \( X = 1 \) where it is reflected, it first travels more rapidly (when it is low down) and then more slowly (when it is high up) than it would in the absence of a magnetic field. The resulting group retardation may be greater or less than that for \( Y = 0 \), according to the distribution of \( X \) (electron density) with height.*

Fig. 11.1. Group refractive index \( \mu' \) as a function of \( X (= 4\pi N_e e^2 / \mu_m m_0) \) for the Ordinary ray and different values of \( Y (= \nu_0 / \omega) \) for a direction of propagation making an angle \( 23^\circ 16' \) with the magnetic field. The scale of \( \mu' \) is proportional to \( 1/\mu' \). (From Shinn & Whale, 1952.)

Curves for the Extraordinary wave take a different form according as \( Y < \) or \( > 1 \). When \( Y < 1 \) the reflection level is reached at \( X = 1 - Y \), so it is convenient to plot \( \mu' \) (in the manner previously described) as a function of \( X / (1 - Y) \) for different values of \( Y \). The resulting curves are shown in fig. 11.2. It will be seen that they all lie above the curve for \( Y = 0 \) so that the group delay is always greater than would be calculated in the absence of a magnetic field.

* In fig. 11.6 an example is illustrated where the group path of the Ordinary wave in the presence of a field (shown as a continuous line) exceeds that with no field (broken line) for some frequencies but not for others.
field. The dotted curves are drawn for $\theta = 0$ (the QL approximation), and it will be seen that they agree closely with the curves for $\theta = 23^\circ 16'$. For the Ordinary wave the QL approximation cannot be used; if it could reflection would occur at $X = 1 + Y$ and not at $X = 1$.

When $Y > 1$ the Extraordinary component is reflected at the level where $X = 1 + Y$, so it is convenient to plot $\mu'$ (in the special form previously mentioned) against $X/(1 + Y)$, as in fig. 11.3. The curves are seen to take a more remarkable form, there is a maximum of $\mu'$ for $X < 1 + Y$, and $\mu'$ falls below its value for $Y = 0$ as $X$ approaches the value $1 + Y$. The dashed line intersects the different curves at the place where $X = 1$, and it is seen that the maxima lie near it. The values of $\mu'[X = 1]$ at $X = 1$ are given by the expression

$$\mu'[X = 1] = 1 + 1/(Y^2 \sin^2 \theta).$$

(11.2.1)

* Correspondingly the curve for the Extraordinary wave in fig. 13.6 is everywhere above that for no field, even when it has been shifted to have the same equivalent penetration frequency.

If $Y$ is just greater than unity a wave-group travelling upwards into the ionosphere will undergo considerable retardation near the level where $X = 1$. A consequence of this retardation is discussed in §13.8.

In the calculations of group velocity made in the papers referred to in the Bibliography of this section it has been supposed that there are no collisions between electrons and heavy particles. Usually the collisions are sufficiently infrequent for the refractive index to be represented quite accurately by the expression (2.5.14) in which collisions are neglected, and then the conclusions of this chapter are correct. If, however, the collisions are more frequent more detailed calculation is necessary.
PART III. APPLICATIONS TO THE TERRESTRIAL IONOSPHERE

CHAPTER 12

A MODEL IONOSPHERE WITH NO MAGNETIC FIELD

12.1. Two model ionospheres

Two different models of the ionosphere will be considered in the discussions which follow. In both it will be supposed that the electron density is horizontally stratified. When the penetration of a layer is to be discussed (§ 13.1) it will be supposed that the electron density increases proportionally to height until it reaches a maximum, above which it decreases. This type of layer is illustrated in fig. 13.2.

When attention is to be restricted to waves reflected in the lower part of a layer, not near the peak, a different model will be used (§§ 12.3 and 13.9). Its properties are represented in fig. 12.1. The electron density increases exponentially upwards as indicated by the line CD which shows the angular plasma frequency \( \omega_N \) as a function of height. \( \omega_N = \left[ 4 \pi N e^2 / e_0 m \right]^{1/3} \) is proportional to \( N^{1/3} \), and the model is constructed so that \( \omega_N / 27 = 10^8 \text{sec}^{-1} \) at a height of 120 km., so that, in the absence of collisions or a magnetic field, a vertically incident wave of frequency 1 Mc/s. would be reflected from this height. It is supposed that \( N \) changes with height \( h \) as given by \( N = N_0 \exp(h/H) \), where \( H = 8 \text{ km} \), so that \( \omega_N \propto \exp(h/2H) \). The way in which the collision frequency \( \nu \) decreases with height is shown by the line AB. It is made to represent closely the conclusions of Nicolet (1953). Although this model has been constructed to have some resemblance to the lower side of the E layer it is only meant to be illustrative and must not be taken to imply anything about the actual structure of that layer.

When the effect of the earth's magnetic field is included the behaviour is somewhat complicated; it is discussed in Chapter 13.

The purpose of this chapter is to introduce the discussion by considering what would happen in the absence of the earth's field.

12.2. The penetration frequency

On a simple ray theory if there were no magnetic field, and if there were no collisions between electrons and heavy particles, a wave incident vertically would reach a level where \( \mu = 0 \) and then would be returned to earth. If \( f = \omega / 2\pi \) is the frequency of the wave the condition \( \mu = 0 \) corresponds to \( X = 1 \) or

\[
4\pi N e^2 / e_0 m \omega^2 = 1,
\]

i.e.

\[
N = (\pi e_0 m e^2)^{1/3},
\]

and if the effective charges are free electrons this gives

\[
N = 1.24 \times 10^{-9} f^2.
\]

If the maximum value of \( N \) in the layer were \( N_m \), then the greatest frequency that could be reflected would be given by (12.2.2) with \( N_m \) written for \( N \). This greatest frequency is called the 'critical' or 'penetration' frequency of the layer.

12.3. The height of reflection

In the presence of collisions the refractive index never becomes zero; its smallest value is \( 1 - 1/(1 + Z^2) \) (see (9.2.10)) and is reached when \( X = 2 \). A full-wave treatment is then required to determine the conditions for reflection. A simplified treatment, given in Chapter 17, shows that that reflection will occur with appreciable amplitude from a level near which \( (\lambda_\mu / \mu^2) (d\mu / dh) \gg 1 \), where \( \lambda_\mu \) is the wave-length in free space and \( h \) is the height. Now if \( Z^2 \ll 1 \) near the level where \( X = 1 \), it has been shown (9.4.10) that \( \mu^2 \ll Z \) and is therefore small; moreover, \( d\mu / dX \) has a finite magnitude, so that, unless the electron density varies very slowly with height, the quantity \( (\lambda_\mu / \mu^2) (d\mu / dh) \) is large. There will therefore be considerable reflection from near this level \( (X = 1) \) even when collisions occur, provided

(a) the wave-length \( \lambda_\mu \) is not too small,
(b) the magnitude of \( Z = v/\omega \) is not too great (say \( Z < 0.1 \)),
(c) the gradient of electron density is not too small.
110 MAGNETO-IONIC THEORY

If the phenomenon of penetration of the layer is omitted the circumstances of reflection at different frequencies can be profitably discussed with reference to the model ionosphere illustrated in fig. 12.1. Provided \( Z < 0.1 \) reflection takes place near the level where \( X = 1 \). This level is shown by the line \( CD \), so long as it is above the line \( EF \) where \( Z < 0.1 \), and the possible levels of reflection are shown by the thickened line \( CG \).

For frequencies less than that corresponding to \( G \), \( Z \) is too large, when \( X = 1 \), to allow \( \mu \) to fall to a small value. For a frequency considerably smaller than this \( Z^2 \gg 1 \) at the level where \( X = 1 \) and, as explained in §9.5, the value of \( \mu \) at this level is near unity and \( \chi \) is near zero. The wave therefore travels upwards, unabsorbed and unreflected, until it reaches a level where either \( d\mu/dh \) or \( d\chi/dh \) becomes large, and, as shown in §9.5, this first occurs near where \( X = Z \). If the wave-length is great enough, and the gradient steep enough, reflection occurs near this level, marked \( HJ \) in the figure.

Fig. 12.1. To illustrate reflection of a wave from a model ionosphere in the absence of an imposed magnetic field. The electron density is represented in terms of the angular plasma frequency \( \omega_p \) by \( CD \) and the collision frequency \( \nu \) by \( AB \). The thick lines represent the heights from which waves of different frequency would be reflected. Although this model is roughly like the lower part of the \( E \) layer it is not intended to represent it in detail.

IONOSPHERE WITHOUT MAGNETIC FIELD 111

The condition \( X = Z \) can be written

\[
\omega_X = \omega,
\]

or

\[
4\pi Ne^2/\varepsilon_0 m = \omega. \tag{12.3.1}
\]

It is interesting to discuss this expression in terms of the permittivity and conductivity of the ionosphere. A medium of permittivity \( \varepsilon_0 \) and conductivity* \( \sigma \) behaves predominantly like a dielectric or like a conductor accordingly as \( 4\pi \sigma \) is less or greater than \( \varepsilon_0 \omega \) and the change-over from the one to the other occurs when

\[
4\pi \sigma/\varepsilon_0 = \omega. \tag{12.3.2}
\]

Now it is known that \( N \) electrons per unit volume, making \( \nu \) collisions per unit time, correspond to a conductivity \( \sigma = Ne^2/2im \), so that (12.3.1) can be written

\[
8\pi \sigma/\varepsilon_0 = \omega. \tag{12.3.3}
\]

If it is also remembered that, for the conditions here envisaged, \( \varepsilon \ (\approx \mu^2) \) is approximately equal to unity, then comparison of (12.3.2) and (12.3.3) shows that reflection of low-frequency waves occurs near the place where the properties of the ionosphere change from those of a dielectric (below) to those of a conductor (above).

If, near the level \( X = Z \), the gradients were not very sharp, and if the wave-length \( \lambda_0 \) were not large, there would be little reflection and the main part of the wave would continue to travel upwards. It would, however, soon be considerably reduced in strength because the absorption coefficient \( \chi \) would be large; indeed, the ionosphere would be like a conducting medium with a gradual lower boundary.

12.4. The equivalent height of reflection

On greater frequencies waves are normally reflected near the height \( (h_0) \), where \( X = 1 \). If a pulse, in the form of a limited wave-train, is sent into the ionosphere it will return from this height after a time which depends on its group velocity at different levels. If it would have taken the same time to travel to a height \( h' \) and back

* Note that \( \sigma \) here does not mean the polarisability of the medium as it did in previous chapters.
with the free-space velocity, then $h'$ is said to be 'the equivalent height of reflection'. It is given by the expression

$$h' = \int_0^{h_1} \mu' dh,$$

(12.4.1)

where $\mu'$ is the group refractive index. If the earth's magnetic field is neglected, and if the effect of collisions is small enough, (11.1.4) shows that

$$\mu' = 1/\mu = (1 - X)^{-1}, \quad \text{where } X = (4\pi e^2/\epsilon_0 m\omega^2)N(h)$$

and $N(h)$ represents the vertical distribution of the electron density.

---

**Fig. 12.2.** To illustrate the shapes of some hypothetical layers by plotting electron density ($N$) against height ($h$).

(a) $N = ah$.
(b) $N = bh^a$.
(c) $N = N_0 \left[ \frac{2h}{h_m} - \left( \frac{h}{h_m} \right)^a \right]$.
(d) $N = N_0 \exp \left[ \frac{1}{2} \left( 1 - \omega/\omega_m \right) - \exp \left( -h/H \right) \right]$, which is an equilibrium Chapman layer for vertical illumination ($\chi = 0$). Equilibrium Chapman layers for other values of $\chi$ have shapes which are like this when the scales are suitably altered.

With different analytical functions taken to represent $N(h)$ it is interesting to calculate the height $h_1$ at which $\mu = 0$, and the equivalent height $h'$. The results for some special cases are given below. In examples (a), (b) and (c) the level $h = 0$ is at the bottom of the layer where $N = 0$.

(a) For $N(h) = ah$ (fig. 12.2(a))

$$h' = 2h_1 \propto \omega^2.$$  

(12.4.2)

(b) For $N(h) = bh^a$ (fig. 12.2(b))

$$h' = \frac{1}{2} \pi h_1 \propto \omega.$$  

(12.4.3)

---

(c) For $N(h) = N_0 \left[ \frac{2h}{h_m} - \left( \frac{h}{h_m} \right)^a \right]$ (fig. 12.2(c))

$$h' = \frac{1}{2} h_m \left( \omega/\omega_m \right) \log \left[ \frac{1 + \omega/\omega_m}{1 - \omega/\omega_m} \right]$$

(12.4.5)

and

$$h_1 = h_m \left[ 1 - \left( 1 - \omega^2/\omega_m^2 \right)^{1/2} \right].$$

(12.4.6)

where $\omega_m$ is the penetration frequency.

(d) For a Chapman layer (fig. 12.2(d)), in which

$$N(h) = N_0 \exp \left[ \frac{1}{2} \left( 1 - \omega/\omega_m \right) - \exp \left( -h/H \right) \right]$$

Jaeger (1947) has shown that, if the wave-frequency $\omega$ is less than the penetration frequency $\omega_m$, so that the wave is reflected in the layer, then

$$h_1 = H \left[ \log \left( \frac{1}{2} \sec \chi \right) - 2 \log \gamma \right].$$

(12.4.8)

where $\gamma = \omega \omega_m$ is a function of the ratio $\omega / \omega_m$ tabulated in his paper. He has also shown that

$$h' = h_1 + \frac{1}{2} HP_1[\omega/\omega_m],$$

(12.4.9)

where the function $P_1[\omega/\omega_m]$ is tabulated.

If $\omega > \omega_m$ the layer is penetrated and the excess of the group path over the geometrical path, for a single traversal of the layer, is given by a quantity $\frac{1}{2} HP[\omega \omega_m]$, which is also tabulated.

---

**12.5. The absorption of the wave**

The total absorption suffered by the wave, in its vertical travel to the level where $\mu = 0$ and back, may be expressed in terms of a reflection coefficient $\rho$ defined as the ratio of the received wave-field to the wave-field which would have been received if there were no absorption. The reflection coefficient is given by

$$\rho = \exp \left( -2 \int_{h=0}^{h=1} \kappa dh \right).$$

(12.5.1)

When there is no imposed magnetic field the absorption coefficient $\kappa$ is given by (4.4.8) with $\omega_L = 0$, i.e. by

$$\kappa = \frac{\nu}{2c} \frac{1}{\mu} \frac{4\pi Ne^2}{\omega_m^2 + \nu^2}.$$  

(12.5.2)

* This is a standard expression for a Chapman layer, in which $\chi = \text{sun's zenith angle}$, $H = \text{scale height}$, $N_0 = \text{value of } N \text{ at the peak of the layer when } \chi = 0$.

See Chapman (1931).
which, with (12.5.1), leads to the expression

\[ -\log \rho = \frac{4\pi e^2}{\epsilon_0 c m} \int_{h=-1}^{h=0} \frac{N\nu}{\mu (\omega^2 + \nu^2)} \, dh. \quad (12.5.3) \]

It is first of interest to consider the magnitude of this absorption in the case of an equilibrium Chapman layer formed in an isothermal atmosphere, so that \( N \) is given by (12.4.7) and

\[ \nu = \nu_0 \exp (-h/H), \quad (12.5.4) \]

\( \nu_0 \) being the value of \( \nu \) at the level where \( N \) is a maximum when \( \chi = 0 \). If, as is often the case, \( \omega^2 > \nu^2 \), Jaeger has shown that when the wave is reflected inside the layer, (12.5.3) leads to

\[ -\log \rho = (\nu_0 H/c \sec \chi) F_{1}(\omega/\omega_m), \quad (12.5.5) \]

where \( F_{1}(\omega/\omega_m) \) is a function of \( \omega/\omega_m \) tabulated in his paper. If, on the contrary, the layer is penetrated, so that \( \omega > \omega_m \), then the absorption suffered by a wave in a double passage through it is given by

\[ -\log \rho = (\nu_0 H \cos \chi/c) F[\omega_m/\omega], \quad (12.5.6) \]

or alternatively by

\[ -\log \rho = (\nu_0 H \omega_m/c \omega^2) (\cos \chi)^{\frac{1}{2}} \phi[\omega_m/\omega], \quad (12.5.7) \]

where \( F[\omega_m/\omega] \) and \( \phi[\omega_m/\omega] \) are tabulated. In (12.5.7) \( \omega_m \) is the angular penetration frequency of the layer when \( \chi = 0 \).

The above expressions give the total absorption of the wave; it is next convenient to consider separately the amount of absorption produced at places (a) where \( \mu \neq 1 \) and (b) where \( \mu \) differs considerably from 1. Since appreciable deviation of an obliquely incident wave can occur only if \( \mu \) is appreciably different from unity, the type of absorption associated with (a) is called non-deviative, and that associated with (b) is called deviative.

(a) Non-deviative absorption

When \( \mu \neq 1 \) (12.5.3) shows that the non-deviative absorption \( \rho_{ND} \) is given by

\[ -\log \rho_{ND} = \frac{4\pi e^2}{\epsilon_0 c m} \int_{h=-1}^{h=0} \frac{N\nu}{\omega^2 + \nu^2} \, dh. \quad (12.5.8) \]

Appleton (1937) has evaluated the magnitude of this expression when \( N \) and \( \nu \) have the values given by (12.4.7) and (12.5.4) appropriate to a Chapman region, and he finds

\[ -\log \rho_{ND} = \{4\pi N_0 e^2 H \exp (\frac{1}{4})/mc^2 \omega \} \psi[\omega \sec \chi/2\nu], \quad (12.5.9) \]

where \( \psi[\chi] \) is a function plotted in fig. 2 of his paper. Since it has been supposed that \( \mu \neq 1 \) at all heights this expression holds only when the layer is penetrated, and \( \omega \) is greater than about \( 2\omega_m \). It applies to a double traverse of the whole layer, and gives the same result as (12.5.6) and (12.5.7) when \( \omega/\omega_m \gg 1 \), and \( \omega^2 \gg \nu^2 \).

It can often be assumed that, over the region responsible for most of the absorption, \( \omega^2 \gg \nu^2 \), so that (12.5.8) becomes

\[ -\log \rho_{ND} = \frac{4\pi e^2}{\epsilon_0 c m \omega^2} \int_{h}^{0} N\nu \, dh. \quad (12.5.10) \]

Under these conditions

\[ -\log \rho_{ND} \propto \omega^{-2}, \quad (12.5.11) \]

and this relation is often said to be characteristic of non-deviative absorption.

If (12.5.10) is used to calculate the absorption produced by a double traverse of a Chapman layer, by inserting \( N \) and \( \nu \) from (12.4.7) and (12.5.4) there results

\[ -\log \rho_{ND} = \{2\pi \exp (\frac{1}{4}) \} \frac{4\pi N_0 e^2 H \nu_0}{\epsilon_0 c m} \omega \langle \cos \chi \rangle^{\frac{1}{2}} \]

\[ = 4.13 \frac{4\pi N_0 e^2 H \nu_0}{\epsilon_0 c m} \omega \langle \cos \chi \rangle^{\frac{1}{2}} \quad (12.5.12) \]

\[ = 4.13 \frac{4\pi N_0 e^2 H N_m \nu_m}{\epsilon_0 c m} \omega \langle \cos \chi \rangle^{\frac{1}{2}}, \quad (12.5.13) \]

where \( N_m \) and \( \nu_m \) represent the magnitudes of \( N \) and \( \nu \) at the peak of the layer whereas \( N_0 \) and \( \nu_0 \) represent their magnitude at the peak when \( \chi = 0 \). Equation (12.5.13) suggests that, if all the absorption of an echo from the \( F \) region were caused by non-deviative absorption in a Chapman-like \( E \) region below, then \( \log \rho \) would be proportional to \( \langle \cos \chi \rangle^{\frac{1}{2}} \), and (12.5.14) shows that this is itself proportional to the product of \( N_m \) the maximum electron density, and \( \nu_m \) the collision frequency at the level of the maximum. Experiments are not in agreement with the expectations, and possibilities of absorption in other parts of the ionosphere have been suggested.
(b) Deviative absorption

If absorption occurs in a region where \( \mu \) is different from unity (4.4.10) may often be used to give valuable results. When there is no magnetic field this equation reduces to

\[
\kappa = \left( \frac{v}{2c} \right) \left( \frac{1}{\mu} - \mu \right).
\]

(12.5.15)

If now \( v \) may be considered to be approximately constant over the region of important absorption, \( \log \rho \) may be expressed as follows:

\[
-\log \rho = 2 \int_{1}^{\infty} \frac{\kappa \, dh}{2 \pi^2} = \frac{v}{2c} \left[ \int_{1}^{\infty} \frac{\kappa \, dh}{\mu} - 2 \int_{1}^{\infty} \frac{\kappa \, dh}{\mu} \right].
\]

(12.5.16)

If there is no magnetic field the quantity \( \left( \frac{1}{\mu} \right) \) represents the group refractive index, so that the two terms in the square bracket represent the group path \( (P') \) and the phase path \( (P) \) traversed by the wave. Hence

\[
-\log \rho = \left( \frac{v}{2c} \right) (P' - P).
\]

(12.5.17)

The value of \( \rho \) given by (12.5.17) can be large either because \( v \) is large or because \( (P' - P) \) is large. In the case of 'non-deviative' absorption \( \mu \) and \( \mu' \) are both near unity so that \( (P' - P) \) is small; the absorption is nevertheless large because \( v \) is large. Equation (12.5.17) is then not very useful.

Under other circumstances, however, \( v \) is not very large and \( \rho \) assumes important measurable values only when \( (P' - P) \) is large. This implies that \( \mu' \) and \( \mu \) depart appreciably from unity, so that considerable refraction occurs for waves incident obliquely, and this type of absorption is called 'deviative absorption'. It is particularly important when a wave is nearly penetrating a region, for then there is a comparatively long path over which \( \mu \) is small and \( \mu' \) large, so that \( (P' - P) \) may be great. Measurements of the change in \( (P' - P) \) associated with a change of \( \log \rho \), as the layer is penetrated, have been used, in conjunction with (12.5.17), to determine the magnitude of \( v \) at the appropriate height.

13. Form of the \( h' (f) \) curves over a wide range of frequency

When a magnetic field is imposed a wave sent vertically into the ionosphere is split into two characteristic waves which travel independently. If the vertical changes in the ionosphere are sufficiently gradual the characteristic waves maintain their independence and the polarisation of each gradually alters so that at any level it is what it would be in a homogeneous ionosphere with the con-

stution appropriate to that level. If there were no collisions and if the imposed magnetic field were not exactly vertical \( \mu \) would change with \( X \) as shown by the curves of fig. 13.1. As the wave ascended \( X \) would increase and reflection of the two components would occur at the places where \( \mu \) first became zero, i.e. where \( X = 1 \) and \( X = 1 - Y \) for \( Y < 1 \) and \( X = 1 \) and \( X = 1 + Y \) for \( Y > 1 \).

Let us now consider what would happen in a simple model ionosphere of the type illustrated in fig. 13.2, in which the electron density \( N \), above a height \( h_0 \), was proportional to \( h - h_0 \) until it reached a sharp maximum value \( N_m \) at a height \( h_m \). It is of interest to plot curves to represent the heights at which \( X \) becomes equal to \( 1 - Y, 1 \), and \( 1 + Y \) for a series of different radio frequencies, and for this purpose it is convenient to normalise all frequencies and express them as multiples of the gyro-frequency \( \omega_H \). In fig. 13.3
MAGNETO-IONIC THEORY

The abscissae represent $\omega^2/\omega_H^2 (= 1/Y^2)$ and are proportional to the square of the radio frequency; the ordinates represent $\omega_N^2/\omega_H^2$ and are proportional to the electron densities $N$, since

$$\omega_N^2 = 4\pi Ne^2/e_m.$$

The ordinates also represent heights for the model ionosphere of fig. 13.2.

The Ordinary wave is reflected at a height where $X = 1$ or $\omega = \omega_N$. This is represented by the line $DKHE$ on the figure. The

![Diagram of electron layer](image)

**Fig. 13.2.** Simple model electron layer.

Extraordinary wave is reflected at one or other of the heights where $X = 1 \pm Y$ or $\omega_N^2/\omega_H^2 = \omega^2/\omega_H^2 \pm \omega/\omega_H$,

and these are represented by the curves $DLIC$ and $BJFA$. The figure can be used as follows to sketch the $h'(f)$ curves which show the equivalent height $h'$ as a function of the frequency $f$.

First suppose that the maximum electron density in the layer is reached at a height $h_m$ corresponding to the line $(a)$ in fig. 13.3, and consider what would happen as the radio-wave frequency was reduced from a great value. Reflection of the Extraordinary wave would first occur from the height $h_m$ at a frequency corresponding to the point $A$, and as the frequency was reduced the height of reflection would decrease until it reached the bottom of the layer ($h_0$) at the gyro-frequency ($B$). If the frequency were reduced to

still smaller values reflection of the Extraordinary wave would occur at heights represented by $CD$. Meanwhile the Ordinary wave would start to be reflected from the height $h_m$ at the frequency corresponding to $E$, and its height of reflection would fall steadily till it reached the level $h_0$ at zero frequency.

In fig. 13.3 the abscissae represent the square of the wave-frequency ($\omega^2 = 4\pi^2 f^2$) and the ordinates represent the height of

![Diagram of curves](image)

**Fig. 13.3.** Normalised curves for use in discussion of heights of reflection. $\omega_N^2/\omega_H^2$ is proportional to $N$, and in the model layer of fig. 13.2 it is proportional to height. The lines $aa$, $bb$ and $cc$ represent different possible levels $h_m$ (see fig. 13.2). The letters on the curves correspond to those of figs. 13.4 and 13.5.

The conclusions of the previous paragraph are represented somewhat differently in fig. 13.4 $(a)$, in which the abscissae represent the wave-frequency ($f$) and the ordinates represent either the real height ($h$) or the equivalent height ($h'$) of reflection. The curves are schematic only and are not meant to correspond quantitatively to those of fig. 13.3. The solid line represents the Ordinary wave and the broken line the Extraordinary. The curves drawn lightly represent the real height of reflection, and correspond to the deductions of the previous paragraph, the letters indicating corresponding points in figs. 13.3 and 13.4 $(a)$.

The curves drawn heavily include the effects of group retardation, and represent the equivalent height ($h'$) of reflection. In interpreting these curves the following facts, mentioned in Chapter 11, must be borne in mind:
For frequencies not too near the gyro-frequency the group refractive index \( \mu' \) reaches large values only when the electron density is near the critical value required to reflect the wave, so that \( X \) is near 1 (Ordinary wave) or \( 1 \pm Y \) (Extraordinary) (see figs. 11.1, 11.2, 11.3 with \( Y \) not too near unity). When reflection occurs from near the maximum of a layer there is a comparatively long path over which \( X \approx 1 \) or \( (1 \pm Y) \) and the integrated group retardation is large. The increased group retardation near \( A \) and \( E \) arise from this cause. It is labelled \( P \) to indicate its association with the penetration frequency.

The form of the \( h'(f) \) curve depends on the maximum ionisation density of the layer. If, instead of being represented by the line \( (a) \) in fig. 13.3, it is represented by the lines \( (b) \) or \( (c) \), then the \( h(f) \) and \( h'(f) \) curves will have the forms sketched in figs. 13.4(b) and (c). Here, again, an effect of group retardation associated with penetration of the layer is labelled \( P \) and one associated with the gyro-frequency is labelled \( G \). The other letters on the curves of fig. 13.4 correspond to those on fig. 13.3.

If there were two layers, one above the other, the form of the \( h'(f) \) curve would depend on whether their maximum electron densities corresponded to the lines \( a, b \) or \( c \) of fig. 13.3. A series of possibilities is illustrated in fig. 13.5. The letters on fig. 13.5 correspond to those on fig. 13.3; the upper and lower sets of letters correspond to the upper and lower layers respectively. Further consideration will be given in §13.8 to the precise form of these curves near the gyro-frequency.

**13.2. Magneto-ionic splitting**

It usually happens that the highest frequencies at which the two components are reflected from any layer are both greater than the gyro-frequency, so that the high-frequency end of the \( h'(f) \) curve is like fig. 13.4(a) or (b). If the layer has a maximum electron density \( N_m \) and if \( \omega_m^2 = 4 m N_m e^2 / \epsilon_0 m \) (so that \( \omega_m \) would be the angular penetration frequency if there were no magnetic field), then the angular penetration frequencies \( \omega_{\text{ord.}} \) and \( \omega_{\text{ext.}} \) for the two components satisfy the relations

\[
X = 1 \text{ (Ordinary)}, \quad \omega_{\text{ord.}} = \omega_m,
\]

\[
X = 1 - Y \text{ (Extraordinary)}, \quad \omega_m^2 = \omega_{\text{ext.}}^2 (1 - \omega_H / \omega_{\text{ext.}}),
\]
MAGNETO-IONIC THEORY

which give
\[ \omega^2_{\text{ext}} - \omega^2_{\text{ord}} = \omega_{\text{ext}} \omega_{H}. \]  
(13.2.1)

If \( \omega_{\text{ext}} \) is greater than about \( 5 \omega_{H} \) then \( \omega_{\text{ext}} \approx \frac{1}{2} (\omega_{\text{ord}} + \omega_{\text{ext}}) \) and (13.2.1) can be written
\[ \omega_{\text{ext}} - \omega_{\text{ord}} \approx \frac{1}{2} \omega_{H}. \]  
(13.2.2)

\[ \begin{align*}
\text{Number} & \quad \text{Top layer type} & \quad \text{Bottom layer type} \\
(i) & \quad \text{L} & \quad \text{I} \\
(ii) & \quad \text{L} & \quad \text{I} \\
(iii) & \quad \text{a} & \quad \text{a} \\
(iv) & \quad \text{b} & \quad \text{b} \\
(v) & \quad \text{a} & \quad \text{a} \\
(vi) & \quad \text{a} & \quad \text{a} \\
\end{align*} \]

Fig. 13.5. The possible types of \( k(f) \) curve for two superimposed layers. The type of each layer is indicated by the letters \( a, b, c \), which correspond to the lines (aa), (bb) and (cc) on fig. 13.3. The frequencies are labelled with capital letters for the top and bottom layers separately, and these letters correspond to those on fig. 13.3.

In the absence of collisions it would therefore be expected that the experimental curve would be split into two by the earth's magnetic field, and that, when the penetration frequencies were not too small, there would be a constant difference between them. This simple expectation is modified if collisions are present, as shown in the next section.

13.3. The effect of collisions on magneto-ionic splitting

When account is taken of the collisions between electrons and heavy particles it is first necessary to decide whether the QT or the QL approximation is appropriate. At the level where \( X = 1 \) the QT approximation will be valid if \( v < \frac{1}{3} \omega_{e} \) and then \( \mu \) for the ordinary wave will decrease rapidly to a low value near this level just as though no field were present. Considerations like those of §12.3 show that, although, in the presence of collisions, \( \mu \) does not reach zero at \( X = 1 \) there will nevertheless be marked reflection near that level if \( Z < c - 1 \).

If, however, \( v > 3 \omega_{e} \) at the level where \( X = 1 \) the QL approximation must be used. For \( Y < 1 \) the two refractive indices now approach zero at the levels where \( X = 1 - Y \) and \( X = 1 + Y \), and reflection will occur near these two places.

If, therefore, \( v < \frac{1}{3} \omega_{e} \) at the level where \( X = 1 \), the two waves are reflected near \( X = 1 \), and \( X = 1 - Y \) (for \( Y < 1 \)) or \( X = 1 + Y \) (for \( Y > 1 \)), as they would be in the absence of collisions, and (13.2.2) represents the relation between the two penetration frequencies. If, however, \( v > 3 \omega_{e} \) at the level where \( X = 1 \), and if \( Y < 1 \), the two waves are reflected near \( X = 1 - Y \) and \( X = 1 + Y \) and the penetration frequencies are given by
\[ \begin{align*}
\omega_{m}^2 &= \omega_{\text{ord}}^2 (1 + \omega_{H}/\omega_{\text{ord}}), \\
\omega_{m}^2 &= \omega_{\text{ext}}^2 (1 - \omega_{H}/\omega_{\text{ext}}),
\end{align*} \]
so that
\[ \omega_{\text{ext}}^2 - \omega_{\text{ord}}^2 = \omega_{H}(\omega_{\text{ext}} + \omega_{\text{ord}}), \]
or
\[ \omega_{\text{ext}} - \omega_{\text{ord}} = \omega_{H}. \]  
(13.3.1)

When \( Y < 1 \) therefore, there would still be two penetration frequencies, but they would be separated by \( \omega_{H} \) which is twice the separation they would have if \( v < \frac{1}{3} \omega_{e} \).

The difference between (13.2.2) and (13.3.1) has been used to show that \( \mu^2 \leq \omega_{e}^2 \) at the level where \( X = 1 \), for waves of frequency about 5 Mc./s.
It may be noted that, if \( Y > 1 \), and if in addition \( \nu > \omega_k \) at the level where \( X = 1 \), then only one wave is reflected, at the level where \( X = 1 + Y \).

If it is assumed that \( \nu^2 \ll \omega_k^2 \) at the level where \( X = 1 \) then insertion of the observed values of \( \omega_{\text{ord.}} \) and \( \omega_{\text{ext.}} \) into (13.2.1) might be expected to yield an accurate value of \( \omega_H \) and hence of the imposed magnetic field \( H \) at the height of reflection. Unfortunately, another phenomenon may complicate the situation and invalidate this method of determining \( H \). This arises because, in the non-isotropic ionosphere, the wave-packets constituting the pulses do not necessarily ascend vertically to the same place, but the two magneto-ionic components are deviated sideways, by different amounts, so that they reach different places, where the maximum electron densities may be different. The details of this sideways deviation of a wave-packet will not be discussed here, but it may be mentioned that values of the two penetration frequencies observed simultaneously in high latitudes appear to be consistent with the supposition that their points of reflection from the \( F \) layer are separated horizontally by about 60 km.

### 13.4 Shapes of the \( h'(f) \) curves for high frequencies

Equation (11.1.4) shows that, in the absence of a magnetic field, the group refractive index \( \mu' = 1/\mu \), and from this result expressions were deduced in § 12.4 to indicate the shapes of the \( h'(f) \) curves for ionospheric layers with different profiles. When a magnetic field is present the expressions for the group refractive indices of the two magneto-ionic components are much more complicated and numerical calculations have to be used.

It is first of interest to consider the shape of the \( h'(f) \) curve for a wave reflected from a model ionosphere of the type which consists of a lower layer in the form of a slab of uniform electron density, and a sharply bounded upper layer which reflects waves of all frequencies from the same height. The group retardation in passing through a slab of unit thickness has been calculated by Shinn & Whale (1952) as a function of frequency, for a slab in which \( \theta = 23^\circ 16' \), \( f_H = 1.2 \text{ Mc./s.} \), and \( f_N = 2 \text{ Mc./s.} \) is the penetration frequency for the Ordinary wave. The \( h'(f) \) curves are shown in fig. 13.6. The curve for the Extraordinary wave is, of course, displaced from that for Ordinary, and in order to compare their shapes, it is also shown shifted so that the penetration frequencies coincide. The broken curve has been calculated for \( Y = 0 \) so that it would apply if there were no imposed field. Although the penetration frequency of the Ordinary wave is not altered by the magnetic field, whereas that of the Extraordinary wave is, it is noticeable that the shape of the two \( h'(f) \) curves are altered to about the same extent, but in different senses. In Shinn & Whale's paper it is shown that experimental \( h'(f) \) curves are distorted in just this way.

One important consequence arises from the way in which the shape of the \( h'(f) \) curve depends on the presence of the earth's field. Attempts have been made to fit the observed curves for the Ordinary wave to the equation (12.4.5) predicted for a parabolic region in the absence of a field, and then to deduce the semi-thickness \( h_m \) and the height of the parabolic layer. Shinn & Whale showed that if this procedure is adopted when a magnetic field is present there may be a considerable error in the value deduced for \( h_m \), but that the value deduced for the height of the layer is reasonably accurate.
126 MAGNETO-IONIC THEORY

The error in the value deduced for $h_m$ depends on (a) the penetration frequency $f_N$, (b) the range of frequencies $f$ used in fitting the experimental curve to (12.4.5) and (c) the angle of dip and the magnitude of the imposed magnetic field.

13.5. Triple splitting and the Z-trace

At places where the angle of dip is fairly steep it is frequently found that the penetration of a layer is accompanied by triple splitting of the $k'(f)$ trace, as shown in the sketch of fig. 13.7. The portions labelled $O$ and $X$ correspond to the Ordinary and Extraordinary waves which are normally observed, and the third portion

![Fig. 13.7. To illustrate the phenomenon of triple splitting.](image)

is called the Z-trace. It is found that the separation between the penetration frequencies of the Z- and X-waves is given by

$$f_X - f_Z = f_H,$$

(13.5.1)

and that the Z-wave has the same polarisation as the Ordinary wave.

Two different explanations of triple splitting have been given. In one (Eckersley, 1950; Rydbeck, 1950, 1951) it is supposed that the third echo arises by a process of coupling, which is most completely described by a full-wave treatment, outlined in Chapter 17. It can, however, be understood simply in the following way.

If the angle between the wave-normal and the earth’s magnetic field is not too great the $\mu(X)$ and $\chi(X)$ curves for the Ordinary and the Extraordinary waves will be like those of fig. 13.8; the Ordinary wave will be reflected near C, where $X = 1$, and the Extraordinary near G, where $X = 1 - Y$. Near the place where $X = 1$, however, both characteristic waves can be propagated, the Ordinary as $\chi$, and hence $R$, of the Ordinary and Extraordinary waves are not very different. When, therefore, an Ordinary wave, travelling into a medium in which $X$ increases gradually, reaches a place corresponding to the shaded region it must change its polarisation rapidly if it is to remain a characteristic Ordinary wave, but, by a smaller change of polarisation it could be propagated as an Extraordinary wave. The result is that some of the energy does, in fact, become transferred to the Extraordinary branch $DEF$ of the curve, and travels upwards to be reflected at $F$, where $X = 1 + Y$. It is this reflection which is responsible for the Z-echo. On the return journey

![Fig. 13.8. The explanation of triple splitting according to Eckersley (1950) and Rydbeck (1950, 1951).](image)
coupling occurs again and some of the energy returns to earth by way of the mode $AB$. In the coupling region, shown shaded between $B$ and $E$, there is little change of polarisation, and the wave behaves much as though it had travelled with left-handed polarisation almost to the reflection point. On emergence from the ionosphere it has left-handed polarisation, just like the ‘Ordinary’ wave which is reflected from the level where $X = 1$.

$$X = 1 + Y$$

Fig. 13.9. The explanation of triple splitting according to Ellis (1953a, b, 1956).

Ellis (1953a, b, 1956) has pointed out that the collision frequency must be fairly great for the ‘coupling’ explanation to be satisfactory except near the poles where the propagation is nearly along the magnetic field. He does not consider that it would be great enough in the $F$ region to explain triple splitting of echoes from that region at moderate latitudes, and he has therefore suggested an alternative explanation, which is illustrated in fig. 13.9. A pulse incident on the ionosphere is split into Ordinary and Extraordinary waves, and the vertically incident Extraordinary wave is reflected at the level where $X = 1 - Y$. If the frequency is a few Mc./s. the $QT$ approximation is valid for the vertically incident Ordinary wave when it reaches

the level where $X = 1$ and it is reflected there. The $Z$-trace is caused by a part of the Ordinary-wave pulse obliquely incident on the ionosphere at an angle $i$. This pulse is refracted as it penetrates and at a level where the refractive index is $\mu$ the wave-normal makes an angle $\phi$ with the vertical given by

$$\sin i = \mu \sin \phi.$$  \hspace{1cm} (13.5.2)

Now it can happen, for a certain value of $i$, that at the level where $X = 1$ the wave-normal direction $\phi$ coincides with the direction $(\phi_H)$ of the earth's magnetic field, so that, even with a small value of $\nu$, the $QL$ approximation holds and $\mu$ does not reach a small value. The wave is therefore not reflected at this level, as it would be for vertical incidence. It continues to travel, until at some greater height it reaches a plane where $X = 1 + Y$. Here $\mu$ becomes small and it is reflected.* Because the reflection is oblique the wave will return to the level $X = 1$ in a direction which is not parallel to the earth's field, so that the $QL$ approximation will no longer apply there. $\mu$ will be small, the absorption large, and the wave will not return to earth. If, however, there are irregularities in the electron distribution near the level where $X = 1 + Y$, where $\mu$ is small, there will be strong† backwards scattering and an appreciable part of the energy will return along the incident path. On the return journey the values of $\mu$ will be the same as on the upward one and, in particular, there will be no difficulty in the wave's penetrating the level where $X = 1$.

The three different waves are returned from levels given by the following expressions:

Extraordinary (vertical)  
$$X = 1 - Y,$$ \hspace{1cm} (13.5.3)

Ordinary (vertical)  
$$X = 1,$$ \hspace{1cm} (13.5.4)

$Z$-component (Ordinary oblique)  
$$X = 1 + Y.$$ \hspace{1cm} (13.5.5)

If the ionosphere is in the form of a layer with a maximum electron density $N_m$, and if the wave-frequency is increased, the three waves

* Since the wave is incident at an angle $i$ it is reflected where $\mu = \sin i$, but in the case of interest $i$ is small (e.g. 10°), and it is sufficiently accurate to suppose that $\mu = 0$.

† More recent work (Pitteway 1958) seems to indicate, however, that backscattering would not be accentuated at a level where $\mu$ was small.

10
penetrate the layer in turn at three different frequencies given, in terms of \( \omega_m^2 = 4\pi N_m^2 e_0 m \) by

- **Extraordinary** \( \omega_m^3 = \omega_{\text{ext}}^2 (1 - \omega_H / \omega_{\text{ext}}) \), (13.5.6)
- **Ordinary** \( \omega_m^3 = \omega_{\text{ord}}^2 \), (13.5.7)
- **Z-wave** \( \omega_m^3 = \omega_Z^2 (1 + \omega_H / \omega_Z) \). (13.5.8)

Equations (13.5.6) and (13.5.7) were considered in §13.2 and represent the usual conditions for magneto-ionic splitting. A combination of (13.5.6) and (13.5.8) shows that

\[ \omega_{\text{ext}} - \omega_Z = \omega_H, \]

so that the difference of the penetration frequencies is equal to the gyro-frequency. It has already been mentioned that observation agrees with this expectation, and that the observed polarisation of the Z-wave is the same as that of the Ordinary, as suggested by the theory.

The suggested explanation can be further tested by comparing the calculated and observed magnitudes of the angle of incidence \( i \) of the Z-wave. If the Ordinary wave travelled throughout with its wave-normal direction along the field direction then \( \mu^2 \) would vary with \( X \) as shown in fig. 13.10 appropriate to longitudinal propagation. It is clear from the simple geometry of this figure that, at \( X = 1 \), \( \mu^2 \) would be given by

\[ \mu^2 = Y / (1 + Y), \] (13.5.9)

![Figure 13.10](image)

Fig. 13.10. To show that \( \mu^2 = Y / (1 + Y) \), when \( X = 1 \), for longitudinal propagation.

Now (13.5.2) relates \( \mu \) and \( \phi \) at any level to the angle of incidence \( \theta \); but at \( X = 1 \), \( \phi = \phi_H \), the angle of the earth's field, so that, combining (13.5.2) and (13.5.9),

\[ \sin^2 \phi = \left( \frac{Y}{(1 + Y)} \right) \sin^2 \phi_H. \] (13.5.10)

Equation (13.5.10) gives the expected angle \( \phi \) of arrival of the Z-trace. It has been shown to give good agreement with the results of experiments.

Waves scattered back from the level where \( X = 1 + Y \) in directions somewhat different from the direction of incidence will not be travelling along the earth's field when they reach the level where \( X = 1 \) and will not be transmitted. Corresponding to a single radiating point on the ground there will be an effective 'hole', at the level where \( X = 1 \), through which the waves can pass upwards or downwards, and waves which do not pass through this 'hole' will not reach the ground. The Z-wave will thus be observable only within a limited distance of the radiating point. Experiment has confirmed this expected result.

### 13.6. Non-deviative absorption, QL approximation

We now consider the absorption of the wave, and, as in §12.5, we shall make the convenient division into non-deviative and deviative absorption. By definition non-deviative absorption occurs at a place where the refractive index is not small, i.e. not near the place where \( X = 1 \) for the Ordinary wave, or \( X = 1 + Y \) for the Extraordinary. The discussion of §8.4 shows that, under these conditions, it will generally happen that the QL approximation is valid, so that the absorption coefficient is given by (4.4.8)

\[ \kappa = \frac{1}{2} \frac{4\pi e^2}{\varepsilon \mu m} \frac{N}{(\omega + \omega_L)^2 + \mu^2}. \] (13.6.1)

It often happens that most of the absorption occurs at levels where \( \mu^2 \ll (\omega + \omega_L)^2 \) so that, from (13.6.1), \( \kappa \propto (\omega + \omega_L)^{-2} \). Under these conditions the reflection coefficient \( \rho \) will obey the relation

\[ -\log \rho \propto (\omega + \omega_L)^{-2}. \] (13.6.2)

Under the same conditions (13.6.1) is equivalent to

\[ \kappa = (\nu / 2\varepsilon) \frac{1}{(1/\mu)} \left\{ X / (1 + Y)^2 \right\}. \] (13.6.3)
This expression has often been used to test whether the absorption of the ordinary wave is of the ‘non-deviating’ type in any particular case. It will be noticed that the absorption coefficient $\kappa$ becomes large near the frequency $\omega_L$, but this does not necessarily imply a large total absorption because the penetration into the layer then becomes small.

13.7. Deviative absorption

(a) Ordinary wave, $QT$ approximation

In this section deviative absorption will first be considered, and reference should be made to § 12.6 where the corresponding problem is discussed for the no-field case. Evidence from the difference of the two penetration frequencies often shows that the $QT$ approximation is valid at the level $X = 1$, where the Ordinary wave is reflected. The results of (9.7.13) then show that, near this level

$$\chi = \frac{1}{2}Z(\mu' - \mu \csc^2 \theta), \quad (13.7.1)$$

and if a calculation like that of § 12.5 is performed there results the expression

$$-\log \rho_D = (v/2c)(P' - P \csc^2 \theta), \quad (13.7.2)$$

which should be used instead of (12.5.17) when there is an imposed magnetic field.

(b) Extraordinary wave, $QL$ approximation

If the collision frequency is small enough the Extraordinary wave will be reflected near one of the levels where $X = 1 \pm Y_L$, and it is of interest to consider the magnitude of the deviative absorption near these heights. It was shown in § 8.3 that, if $\theta < 40^\circ$, the $QL$ approximation is valid when $X = 1 \pm Y_L$, and then

$$\mu' = 1 - X/(1 \pm Y_L), \quad (13.7.3)$$

and $\kappa$ is given by (13.6.3).

It is shown in the Appendix that, under these conditions,

$$\kappa = (v/2c)(\mu' - \mu)/(1 - \frac{1}{2} Y_L), \quad (13.7.4)$$

so that

$$-\log \rho_D = \int \kappa dh = (v/2c)(P' - P)/(1 - \frac{1}{2} Y_L). \quad (13.7.5)$$

This expression gives the deviative absorption of the Extraordinary wave if $\theta < 40^\circ$, so that the $QL$ approximation is valid at the reflection level.

13.8. Gyro-frequency splitting

The discussion of § 13.1 and the curves of figs. 13.4 and 13.5 indicate that a splitting of the echo-trace should be observed near the gyro-frequency. Three different sets of workers have observed this type of splitting, but they do not agree on the observed facts, or on the explanation of the observations. It will therefore be profitable to discuss their results in some detail.

The observations of Martyn & Munro (1938 a, b, 1939) (MM) were made in Australia, those of Booker & Berkner (1938 a, b) (BB) in America, and those of Farmer & Ratcliffe (1935), later discussed by Appleton, Farmer & Ratcliffe (1938) (AFR), in England. All worked at night and observed a splitting of the $F$ region echoes, with the upper component increasing in height as the frequency increased towards the gyro-frequency. There is disagreement about the frequency at which the echo-path approached infinity; MM consider that the frequency of infinite delay was $\omega_H \cos \theta$ ($= \omega_L$); BB think it was sufficiently smaller than either $\omega_H$ or $\omega_L$, and their conclusion is based on a good published record; AFR did not observe the frequency of infinite delay. There is also disagreement about the polarisation. AFR received on a circularly polarised aerial and found both echoes to be fairly accurately circularly polarised, the upper with right-handed and the lower with left-handed sense of rotation. MM found the upper echo to be of mixed polarisation, about half the time linear, and half the time elliptical with left-handed sense. They worked in the Southern Hemisphere, and this left-handed polarisation corresponds to right-handed in the Northern Hemisphere. BB did not observe polarisation.

(a) The explanation of Appleton, Farmer and Ratcliffe

AFR consider that their observations could be explained if, at the time of the experiment, the penetration frequency of the $E$ layer corresponded to the line (c) in fig. 13.3 and that of the $F$ layer to either (a) or (b), so that the $h'(f)$ curve would be as represented either in fig. 13.5 (ii) or 13.5 (ii). If fig. 13.5 (ii) represents the facts then the frequency $I$ of great delay could differ from both $f_H$ and $f_L$, as was observed by BB. The $F$ layer would be penetrated twice by the Extraordinary wave, at the two frequencies marked $I$ and $F$, ...
and since fig. 13.3 shows that these frequencies correspond to the conditions \(X = 1 + Y\) their difference should be given by \(f_H\) (see (13.3.1)). Reference to the record published by BB shows, however, that the difference between the two frequencies is considerably greater than \(f_H\), so that the explanation of the splitting cannot be as represented in fig. 13.5 (ii).

If the facts were represented by fig. 13.5 (iii), then the gyro-splitting would occur at the frequency \(C\) which does not represent the penetration of any layer, it simply represents the highest frequency which can be reflected at the level where \(X = 1 + Y\). To calculate the shape of the \(k'(f)\) curve we write

\[
\frac{d\mu}{dY} \approx \frac{d\mu}{dX} \frac{dX}{dY} = \frac{d\mu}{dX} \frac{dX}{dY}.
\]

and evaluate the integral for different frequencies. When a layer is penetrated the region near the ionisation maximum, where \(d\mu/dX = \infty\), plays a large part in determining the value of this integral, and it is responsible for the marked group retardation accompanying penetration. Under the circumstances here considered, however, the layer is not penetrated, and another reason must be found for the increased group path near the frequency \(C\). The cause can be understood from the curve labelled \(Y = 1\frac{1}{2}\) in fig. 11.3. This curve shows that, when \(Y\) is just greater than 1, \(\mu'\) has large values for values of \(X\) less than about \(\frac{1}{2}(1 + Y)\), so that if \(d\mu'\) is evaluated up to the level where \(X = 1 + Y\), the major contributions come from the levels where \(0 < X < \frac{1}{2}(1 + Y)\). As the gyro-frequency is approached the values of \(\mu'\) become greater and \(k'\) increases towards infinity.

The region responsible for the large group retardation is mainly low down in the ionosphere where \(X < 1 + Y_L\), and may well be in the \(E\) layer even when the wave is reflected from the \(F\) layer. BB have pointed out that the \(k'(f)\) records lend support to this suggestion, and that the so-called '\(M\)-echoes' which have traversed a path such as that shown in fig. 13.11 experience only the same retardation as echoes once reflected from the \(F\) layer.

If the explanation of AFR were correct the frequency of greatest delay would be expected to be \(f_H\) if there were no collisions. If the electrons collided with heavy particles it might happen that \(v \gg \omega_e\) and near the level where \(X = 1\), so that the \(QL\) approximation would be valid. The frequency of greatest delay would then be \(f_H \cos \theta (= f_L)\). Presumably intermediate values of the collision frequency would give rise to intermediate values for the frequency of infinite delay.

(b) The explanation of Martyn and Munro

MM reject the simple explanation of AFR because they did not find the upper echo to be circularly polarised. They suggest that the upper echo arises by a process of coupling between two of the characteristic waves, as illustrated in fig. 13.12. The Ordinary wave is normally strongly reflected near the level \(A\) where \(X = 1\), at which the refractive index would be zero if there were no collisions. Above that level there is a 'barrier region' (\(\mu\) is imaginary with no collisions, absorption large with collisions) and then at

\[
X = \frac{(1 - Y^2)}{(1 - Y_L^2)}
\]

there is a sudden and great increase in \(\mu\). The suggestion is that an appreciable part of the energy crosses this barrier and is reflected near the level \(B\) where \(\mu\) becomes large. Both echoes from \(A\) and \(B\) represent the Ordinary wave, and it might be expected that they would both have the circular polarisation with right-handed sense

* It is mentioned in §13.9 that Budden (1955b) considers that reflection will not occur at a level where \(\mu\) becomes large.
appropriate to it in the Southern Hemisphere. Since, however, the polarisation changes rapidly as the wave passes the level where \( X = 1 \), it is suggested that some left-handed component is introduced into the upper echo, reflected at \( B \), and that this accounts for the mixed polarisation observed by MM. In their explanation, the wave which might be expected to be reflected at the level where \( X = 1 + Y \) is assumed to be completely removed by absorption.

\[ X = \frac{1 - Y^2}{1 - \frac{Y}{L}} \]

Fig. 13.12. To illustrate the explanation of splitting near the gyro-frequency given by Martyn & Munro (1938).

If the wave-frequency is increased, \( Y \) and \( Y_L \), which are both less than unity, will increase, and when \( Y_L = 1 \) the electron density corresponding to the point \( B \) will become infinite. This would imply that the frequency of infinite delay was \( f_H \cos \theta \). MM suppose that the wave reflected from \( B \) travels, not vertically, but nearly along the earth's field, so that this frequency is effectively the gyro-frequency \( f_H \).

(c) The explanation of Booker and Berkner

The experimental curves of BB showed a great retardation of the \( F \)-layer echo at a frequency which was about \( 0.9f_H \). They considered that their observations, and their estimate of the magnitude of \( f_H \) at the appropriate height, were good enough to make the observed frequency significantly different from \( f_H \), and they therefore rejected the simple explanation of AFR.

They suggested instead that the results could be explained if the

'IoRT Term' were included in the equations of the magneto-ionic theory. A theory in which this term is included was called by them a Lorentz theory, in contradistinction to a Sellmeier theory which is the one dealt with in the main part of this book. The Lorentz theory is discussed in Chapter 15, which should be referred to in relation to the following discussion.

Goubau (1934,6) showed, from the Lorentz theory, that when the angle \( \theta \) of propagation is less than \( \cos^{-1}(1/\sqrt{3}) \), the Extraordinary wave suffers considerable group retardation for small values of \( X \) when the frequency is such as to give \( Y \) a value just less than a quantity \( Y_1 \) given in terms of \( \theta \) by fig. 15.3 and that the retardation becomes infinite when \( Y = Y_1 \). On this theory the great group retardation occurs low in the ionosphere, just as it does on the Sellmeier theory (see explanation of AFR), but at a different frequency. For the place where the observations were made fig. 15.3 shows \( Y_1 \) to be equal to 1.17. This agreed roughly with the observed value \( f_H/1.1 \) for the frequency of infinite delay, but the accuracy was thought to be such that there was a significant discrepancy.

It was suggested that the discrepancy could be removed if it was supposed that there were an important number of heavy ions, in addition to the electrons, at the level, in the \( E \) region, where the main group retardation was produced. Calculation showed that the heavy ions would have to be about \( 10^4 \) times as numerous as the electrons.

If this explanation is accepted it must be supposed that the Lorentz theory is correct. But other observations, discussed in Chapter 15, suggest that the Sellmeier theory is the correct one, and the present view is that the Lorentz term should not be included in the equations. It thus appears that a proper explanation of the splitting on frequencies just less than the gyro-frequency has still to be found.

13.9. The 'fourth reflection condition'

In the curves of fig. 13.1 the refractive index of one or other of the characteristic waves becomes infinite when \( X \) reaches a certain

- The appropriate \( \mu(X) \) curve then goes to infinity for a comparatively small value of \( X \), as in fig. 15.2(c).
value. It has often been suggested (see Bibliography) that if a wave travelling through the ionosphere reached a level where $X$ has this value it would be reflected, and the conditions which would produce this reflection have been called the fourth reflection conditions, the other three being when $X = (1 - Y)$, 1, and $(1 + Y)$.

It might be objected that, since the value of $X$ which corresponds to an infinity in $\mu$ is always greater than that which corresponds to $\mu = 0$, a wave entering the ionosphere from below would never reach the place where $\mu = \infty$. If, however, the distance between the levels where $\mu = 0$ and $\mu = \infty$ were not too great it has been suggested that some 'leakage' might occur through the 'barrier' (see § 17.1). It has also been suggested that an Ordinary wave with $Y > 1$, which had entered the ionosphere by passing the level where $X = 1$ when the collision frequency was great, might, on emerging from the upper side of the ionosphere, encounter the infinity in $\mu$ as $X$ decreased, and might be reflected.

Budden (1955 b) has used a full-wave analysis to consider these matters in detail, and has concluded that reflection would not be expected near where $\mu = \infty$, and that the fourth reflection condition is, in fact, a misconception. He concludes that the energy is neither reflected nor absorbed near that level but that it is laterally deviated towards the nearer magnetic pole. His argument disposes of another suggestion which has sometimes been made, that because the group velocity becomes zero when $\mu = \infty$ it must follow that reflection will then take place.

13.10. Low and very low frequencies and the lowest ionosphere

So far this chapter has been concerned chiefly with waves of frequency near and above the gyro-frequency and with phenomena concerned with the penetration of ionospheric layers. In this section lower frequencies and the lower parts of the ionosphere will be considered, and for this purpose it will be convenient to suppose that the electron density increases upwards exponentially without limit. The model previously discussed in § 12.1 and fig. 12.1 will again be used, and the argument of that section will be modified to take account of the imposed magnetic field. It will be convenient to extend the discussion to frequencies greater than the gyro-frequency so as to link up with the previous parts of this chapter, but matters of layer penetration will not be discussed.

The model of fig. 12.1 is repeated by fig. 13.13 in which the collision frequency $(\nu)$ and the electron density (proportional to the square of the plasma frequency $\omega'_{p}$) are represented by the lines $AB$ and $CD$ in the same way as before. The frequency labelled $\Omega$ in this figure is important. It is that frequency which makes $X = 1$

\[
X = 1 - Y
\]

\[
\omega'_{p} = \omega'_{p} - \omega_{0}
\]

\[
\omega'_{p} = \omega'_{p} + \omega_{0}
\]

\[
\omega'_{p} = \omega'_{p}
\]

at the level where $\nu = \omega_{c}$. The level at which $\nu = \frac{1}{2} \omega_{c}$ is used to mark in the regions of validity of the QT approximation as explained in § 8.4 and fig. 8.1, and the lines $X = 1 + Y$ and $X = 1 - Y$ are also drawn in. If, then, $\nu$ were nowhere great enough to prevent $\mu$ falling to small values near $X = 1$ or $(1 \pm Y)$ reflections would occur as follows:

(a) So long as the QT approximation holds at $X = 1$ the Ordinary wave will be reflected from that level as indicated by the thick continuous line $CG$. The range of angular frequencies concerned runs from just above $\Omega$ up to infinity.

\[
\nu = \omega_{c}
\]

\[
\omega'_{p} = \omega'_{p} - \omega_{0}
\]

\[
\omega'_{p} = \omega'_{p} + \omega_{0}
\]
(b) For angular frequencies greater than \( \Omega \) the Extraordinary wave will be reflected from the level where \( X = 1 - Y \) (the line \( CLH \)) if \( \nu > \nu_E \) and from the level where \( X = 1 + Y \) (the dashed line \( CI \)) if \( \nu < \nu_E \).

(c) For angular frequencies less than \( \Omega \) the Ordinary wave will be reflected from the level where \( X = 1 + Y \) as represented by the continuous line \( IK \). There has been some confusion about what happens when \( \nu > \nu_E \) at \( X = 1 \) and decreases at higher levels so that there \( \nu < \nu_E \). It has been suggested (Lepchinsky, 1956) that reflection would then occur from near the level where \( \nu = \nu_E \), but this conclusion does not appear to be correct (Landmark & Lied, 1957).

It is now necessary to investigate the way in which collisions can interfere with reflections by preventing \( \mu \) approaching zero. The results of \( \S 10.3 \) are used for this purpose.

If \( Z > \nu \cdot (1 \pm Y) \) then the value of \( \mu \) at the levels \( X = 1 \pm Y \) cannot become small (10.3.6). The lines \( Z = \nu \cdot (1 \pm Y) \) or \( \nu = \nu \cdot (1 \pm Y) \) are shown in the figure and they intersect the curves for \( X = 1 \pm Y \) at the points \( J \) and \( L \) which represent the lowest frequencies capable of being reflected on these two curves. The absence of reflection for frequencies between \( \nu_E \) and that represented by \( L \) is further discussed in \( \S 13.11 \) at (b).

For frequencies much less than that corresponding to \( J \) the discussion of \( \S 9.5 \) is applicable. For the Ordinary wave \( Z^2 \gg 1 \) when \( X' = 1 \) and no reflection is possible at the level where \( X' = 1 \) (i.e. \( X = 1 + Y \)). As the wave ascends, however, \( Z' \) decreases and \( X' \) increases until a level is reached where \( Z' = X' \) (or \( Z = X \)), near which there is a marked change in \( \chi \) (not \( \mu \)) and reflection can take place if this is sufficiently rapid. These changes occur for both the characteristic waves, which are therefore both reflected. This possible reflection level is represented by the line \( MN \).

At frequencies of 16 kc./s, full-wave calculations have been made by Heading & Whipple (1952) for a wave incident on an ionosphere with an exponentially increasing electron density, and in which \( Y \gg Z \). In the model which they considered conditions corresponded to those between the frequencies \( M \) and \( J \) on fig. 13.13.

In the calculations it was convenient to consider separately two levels near which \( X = Z \) (MP in the figure) and \( X = Y \) (MJ in the figure). These they called Regions I and II.

Many full-wave calculations have been made for a frequency near 150 kc./s. (\( \omega = 9.4 \times 10^8 \)), and it has been suggested that phenomena of ‘coupling’* between the Ordinary and Extraordinary wave can play an important part near the level where \( X = 1 \), particularly if, at that level, \( \nu \approx \nu_E \). One effect of the coupling is to produce an echo from the level where \( \nu = \nu_E \) when the frequency is \( \Omega \), so that this level is also the place where \( X = 1 \). This echo has been called a ‘coupling echo’. It has been observed simultaneously with a ‘main echo’ reflected from the level, corresponding to the point \( I \) in fig. 13.13, where \( X = 1 + Y \).

A second effect is concerned with the polarisation of the ‘main echo’ when the frequency is near \( \Omega \). At a frequency slightly greater than \( \Omega \) the polarisation is Extraordinary (dashed line), and at a frequency slightly less it is Ordinary (continuous line). The way in which the refractive indices change with height for the two neighbouring frequencies is illustrated in figs. 13.14d and e, which show how a small change of frequency can result in the reflection of a differently polarised wave from the level where \( X = 1 + Y \). If, now, on a fixed frequency, the ionisation density were to alter, the point \( I \) could move so that at one time it represented a higher frequency and at another time a lower one, and on a simple ray picture the received polarisation would change suddenly from that appropriate to Ordinary to that appropriate to Extraordinary. It would be accompanied by a reversal of the sense of rotation in the returned wave, and a change in the orientation of the ellipse. This phenomenon has been observed and is called the ‘flip-over’ phenomenon.

13.11. \( \mu(h) \) and \( \chi(h) \) curves for different frequencies in the model ionosphere of fig. 13.13

It is instructive to draw curves to show how \( \mu \) and \( \chi \) would vary with height for a series of different frequencies for the model ionosphere of fig. 13.13. The results, which should be compared with fig. 13.13, are shown in fig. 13.14, together with some notes on the next page.

* These phenomena are discussed more fully in \( \S\S 17.3, 17.4 \) and 17.5.
Notes on fig. 13.14

(a) $\omega = 3 \times 10^7$

Reflection of the Ordinary wave occurs near $X = 1$ and of the Extraordinary near $X = 1 - Y$.

(b) $\omega = 10^7$

Collisions are sufficiently frequent to prevent $\mu$ falling to a small value near $X = 1 - Y$, so there is no reflection of the Extraordinary wave there. It is not reflected at $X = 1 + Y$ because it is strongly absorbed at levels just above $X = 1 - Y$. It should be particularly noticed that the absence of reflection at $1 - Y$ is not caused by absorption below $1 - Y$, but by the large magnitude of $\mu$ at $1 - Y$. Becker (1950) has pointed out that these circumstances explain the frequent occurrence of $h'(f)$ curves in which the Ordinary echo from the $E$ region is absent on frequencies less than about 2 Mc/s.

(c) $\omega = 3 \times 10^8$

Reflection of the Ordinary wave occurs near $X = 1$ and of the Extraordinary wave near $X = 1 + Y$.

(d) $\omega = 10^8$

Collisions are sufficiently frequent to prevent $\mu$ falling to zero near $X = 1$, so there is no reflection of the Ordinary wave there. The wave which is reflected from $X = 1 + Y$ is the Extraordinary one since $\nu < \omega_k$ where $X = 1$.

(e) $\omega = 3 \times 10^8$

Since $\nu > \omega_k$ when $X = 1$ the propagation is now $QL$ throughout and the wave which is reflected from $X = 1 + Y$ is the Ordinary one.

(f) $\omega = 3 \times 10^9$

When $X = 1 + Y$ the collisions are so frequent that for each characteristic wave $\mu$ is not allowed to depart much from unity.

(g) $\omega = 3 \times 10^9$

When $X \neq (1 + Y), Z > 0.1(1 + Y)$ and neither value of $\mu$ can depart appreciably from unity. When

$$Z < 0.1(1 + Y), \quad X > 2(1 + Y),$$

and still neither value of $\mu$ can depart appreciably from unity. As the level $X = Z$ is passed, however, $\chi$ for both waves suddenly becomes large and reflection may occur.

In the above discussion it has been supposed that the levels of reflection could be accurately located near the places where $\mu$ or $\chi$

![Graphs and diagrams showing variation of $\mu^2$ and $\chi^2$ with $X$ (or height) for different frequencies in the model ionosphere of fig. 13.13.]

changed rapidly. This simple 'ray' picture is, of course, much oversimplified; it is probably reasonable for frequencies greater than 500 kc/s, doubtfully correct for frequencies between 200 and 500 kc/s, and incorrect below 200 kc/s. On these lower frequencies a 'full-wave' treatment must be used.
CHAPTER 14

THE NATURE OF THE ELECTRONS' MOTIONS IN THE IONOSPHERE

14.1. Magnitudes in the absence of a magnetic field

In Chapters 3 and 5 the microscopic mechanism responsible for refraction and absorption in the ionosphere was discussed in terms of the detailed motions of the electrons and their energies. It is of interest to consider numerical magnitudes of these quantities, appropriate to the terrestrial ionosphere. For the purpose of illustration the E layer will be considered and the following values will be taken as representative:

<table>
<thead>
<tr>
<th>Table 14.1. Approximate magnitudes in the E layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
</tr>
<tr>
<td>Collision frequency</td>
</tr>
<tr>
<td>Mean free path of electrons</td>
</tr>
<tr>
<td>Mean gas-kinetic velocity of electrons</td>
</tr>
<tr>
<td>Mean gas-kinetic energy of electrons</td>
</tr>
<tr>
<td>Gyro-frequency</td>
</tr>
</tbody>
</table>

Now suppose that a sender* radiates a power of 100 kW, and suppose (what is impossible) that this power is radiated isotropically. Then the power flux at the E layer, supposed to be at a height of 100 km., is $8 \times 10^{-7} \text{ W.m}^{-2}$. This corresponds to an electric field, in free space, of about $0.17 \text{ V.m}^{-1}$. This field $(\mathcal{E})$ is independent of the wave-frequency. It sets an electron into oscillation with an amplitude $(x_0)$ given by

$$x_0 = \mathcal{E}e/m\omega^2.$$  \hspace{1cm} (14.1.1)

The maximum velocity $(v_0)$ of the electron in its oscillation is given by

$$v_0 = \mathcal{E}e/m\omega,$$  \hspace{1cm} (14.1.2)

and its mean oscillatory energy $U_0$ by

$$U_0 = \mathcal{E}e^2/4m\omega^3.$$  \hspace{1cm} (14.1.3)

The magnitudes of the quantities $x_0$, $v_0$ and $U_0$ are shown in Table 14.2.

---

* These figures are typical of a high-power broadcasting station.
columns (3), (5) and (7) of table 14.2 for a series of different frequencies. The ‘thermal’, gas-kinetic magnitudes from table 1 are included for comparison.

14.2. Heating of the electrons and ionospheric cross-modulation (imposed field neglected).

As explained in §5.2 the mean ordered oscillatory energy $U_o$ of each electron is converted, at the end of the free path, into disordered ‘heat’ energy, so that the electrons ‘heat up’. If there were no process by which they could ‘cool down’ they would heat up indefinitely so long as the wave was applied. But when the electrons acquire a mean energy $U_e$ greater than the mean thermal energy $U_o$ of the surrounding molecules they give up energy to those molecules at each collision. The average energy $\Delta U$ given up at each collision is thought, from laboratory experiments, to be given by

$$\Delta U = G(U_e - U_o), \quad (14.2.1)$$

where $G \approx 1.3 \times 10^{-3}$.

A state of statistical equilibrium will be reached when the average thermal energy $U_o$ gained in each free path, by the conversion of ordered oscillatory energy at a collision, is equal to the average energy $\Delta U$ lost to the molecules at each collision, so that

$$U_o = G(U_e - U_o). \quad (14.2.2)$$

If a value of $U_o$ appropriate to a high-power broadcasting station ($100$ kW. at $1$ Mc./s.) is inserted from table 14.2 this relation gives

$$5 \times 10^{-18} = 1.3 \times 10^{-3}(U_e - U_o),$$

and insertion of the magnitude of $U_e$ from table 14.1 shows that

$$(U_e - U_o)/U_o = 6.3 \times 10^{-3}.$$

It therefore follows that the mean energy of the electrons exceeds the mean thermal energy of the molecules by about $0.6\%$ when the wave is present, so that the electron ‘temperature’ is also raised by this amount.

It might be thought surprising that the presence of the wave could modify the electron temperature so much, particularly in view of the fact that the mean oscillatory energy of the electron ($5 \times 10^{-18}$ erg) is small compared with its mean thermal energy $(6 \times 10^{-14}$ erg). It must be remembered, however, that the electron is able to get rid of only a small fraction of its excess energy at each collision, and it will therefore continue to ‘heat up’ until its excess energy is such that the small fraction which can be lost is sufficient to balance what it receives from the wave.

If the electrons were heated up by the presence of a wave until their mean energy was $U_1$, and if the wave were then switched off, they would lose their energy $U_o$ according to the relation

$$dU_e/dt = -G(V_e - U_o), \quad (14.2.3)$$

since the energy lost at each collision would be $G(U_e - U_o)$ and there are $\nu$ collisions in unit time. The average energy would thus fall exponentially as given by

$$U_e = U_o + (U_1 - U_o) \exp(-G\nu) \quad (14.2.4)$$

till it was equal to the thermal energy of the molecules. The time constant in this expression is $(G\nu)^{-1}$ and is about $10^{-3}$ sec.

The increase of electron temperature, produced when a wave is absorbed in the ionosphere, is accompanied by an increase in the collision frequency. But it is the collision frequency which is responsible for the absorption of waves, so that when one wave ($A$) is passing through, and heating up, the ionosphere, another wave ($B$) will be more strongly absorbed. This phenomenon has been demonstrated by switching wave $A$ on and off; it is, however, more obviously recognisable if $A$ is sinusoidally modulated so that the absorption of $B$ varies sinusoidally; it then emerges from the ionosphere with the modulation of $A$ imposed upon it. This phenomenon has been called ‘ionospheric cross-modulation’, ‘wave interaction’, or ‘the Luxemberg effect’.

Since the time constant of the heating and cooling of the electrons is $(G\nu)^{-1}$, the phenomenon is most noticeable when the modulation frequency is less than $G\nu$, i.e. less than about $100$ sec.$^{-1}$.

14.3. Magnitudes in the presence of a magnetic field

When a steady magnetic field is present the gas-kinetic motion of the electrons is different. In a free path between collisions the component along the field is unaltered, but across the field the electrons are constrained to move in circles with the angular gyro-
frequency $\omega_H$. They all move round the field in the same sense, and if $v'_\theta$ is the component of gas-kinetic velocity perpendicular to the magnetic field the radius ($r_\theta$) of the circle is

$$r_\theta = \frac{v'_\theta}{\omega_H}.$$  \hspace{1cm} (14.3.1)

Orders of magnitude can be obtained by using the numbers in table 14.1.

If the average value of $v'_\theta$ is taken to be $\frac{1}{2}v_\theta$ and

$$\frac{\omega_H}{2\pi} = 1 \text{ Mc./s},$$

then $r_\theta = 0.6$ cm. The average free path will thus be in the form of a spiral of radius $0.6$ cm., length $100$ cm. and containing

$$\frac{\omega_H}{2\pi r} = 10 \text{ revolutions}.$$

The application of an electro-magnetic wave alters the motion of an electron differently according to the direction of propagation. If the QT approximation is appropriate the electric field in the Ordinary wave is along the direction of the applied magnetic field, and it alters the motions of the spiralling electrons just as it would if no field were present. The Extraordinary wave produces more complicated motions, not easy to describe.

If the $QL$ approximation is valid the fields of the characteristic waves drive the electrons in circles, described with the wave-frequency, in planes perpendicular to the imposed magnetic field. These circles are superimposed on the free circular motion in the gas-kinetic spiral, that of the Extraordinary wave is in the same sense as the free rotation and that of the Ordinary is in the opposite sense.

If the circularly polarised wave has a field strength $\mathcal{E}/\sqrt{2}$, it will produce the same total power flux as a plane-polarised wave with field $\mathcal{E}$. The radius $r_H$, the velocity $v_H$, and the energy $U_H$ in the driven circular orbits are given, for such a wave, by the following expressions which relate them to the corresponding quantities $x_0$, $v_0$ and $U_0$ of (14.1.1), (14.1.2) and (14.1.3), for a linearly polarised wave of the same total power flux. The upper sign corresponds to the Ordinary wave:

$$r_H = \frac{(\mathcal{E}e/m \sqrt{2})/(\omega (1 \pm \omega_H/\omega))}{(x_0/\sqrt{2})/(1 \pm \omega_H/\omega)} = (x_0/\sqrt{2})/(1 \pm \omega_H/\omega),$$  \hspace{1cm} (14.3.2)

$$v_H = \frac{(\mathcal{E}e/m \sqrt{2})/(\omega (1 \pm \omega_H/\omega))}{(v_0/\sqrt{2})/(1 \pm \omega_H/\omega)} = (v_0/\sqrt{2})/(1 \pm \omega_H/\omega),$$  \hspace{1cm} (14.3.3)

$$U_H = \frac{(\mathcal{E}e^2/4m)/(\omega (1 \pm \omega_H/\omega))^2}{(x_0/\sqrt{2})/(1 \pm \omega_H/\omega)} = U_0/(1 \pm \omega_H/\omega)^2.$$  \hspace{1cm} (14.3.4)

![Fig. 14.1](image-url)

The magnitudes of these quantities for the Ordinary wave are given, for different frequencies, in columns 4, 6, 8 of table 14.2, in which, for simplicity, the gyro-frequency has been taken to be 1 Mc./s.

It will be seen that, for the Ordinary wave, and for frequencies greater than the gyro-frequency, the magnitudes are not much different from those of columns (3), (5) and (7) which apply when there is no applied field. For frequencies less than the gyro-frequency, however, the magnetic field limits the excursions of

the electrons so that their velocities and energies no longer reach the relatively large values of columns (3), (5) and (7).

For the Extraordinary wave ($QL$ approximation) (14.3.2), (14.3.3) and (14.3.4) show that the magnitudes of $r_H$, $v_H$ and $U_H$ become great when the wave-frequency approaches the gyro-frequency. The ratio of the Extraordinary to the Ordinary values of the radius $r_H$ or the velocity $v_H$ are given by the ratio

$$\left[(1 + \omega_H/\omega)/(1 - \omega_H/\omega)\right],$$

which is shown as a function of $\omega_H/\omega$ (or $Y$) in fig. 14.1. The
square of this function represents the ratio of corresponding values of the energy.

The circularly polarised component of the Ordinary wave produces a rotation of the electron in the opposite sense to that of the free gas-kinetic rotation, and the Extraordinary wave produces a rotation in the same sense as the free rotation. The combination of the forced circular motions described with the wave-frequency, and the free circular motion, described with the gyro-frequency, therefore produce orbits which would look like those of fig. 14.2 to an observer looking along the magnetic field. In this figure it is supposed that the free orbit is three times as large as the forced orbit at all frequencies.

14.4. Energy relations in the presence of a magnetic field

At the end of each free path the ordered energy is converted into thermal energy, as explained in §5.2, and, so far as the ordered movement is concerned, the electron is left at rest. Just after the collision the ordered oscillation has to be restarted, and since a finite extra ordered velocity cannot be produced instantaneously a 'transient' occurs, which, when added to the 'forced' movement, allows the velocity to build up gradually. The details of this 'transient' when there was no applied field were discussed in §5.2. It is interesting to consider its nature when an imposed field is present.

The matter can be understood by neglecting the gas-kinetic spiral motion and considering how an electron, originally at rest, would move if the circularly polarised wave-field of one of the $QL$ characteristic waves was suddenly applied to it. The 'forced' oscillation would be in a circle with the frequency of the applied wave and with a velocity given by (14.3.3). Because this velocity could not be produced instantaneously a 'transient' would also occur which would, at the start, provide a cancelling velocity equal in magnitude and opposite in direction. This would take the form of a circle described with the free gyro-frequency. If the radii of the forced and the gyro-frequency circles are $r_f$ and $r_g$ respectively, then because the linear velocities in the two circles are the same

$$v = \omega r_f = \omega_H r_g,$$

so that

$$r_f/r_g = \omega_H/\omega = Y. \quad (14.4.1)$$

Fig. 14.2. To illustrate the projected movement of an electron travelling freely between two collisions when it is acted on by a circularly polarised wave-field in the plane of the paper and of angular frequency $\omega$, in the presence of an external magnetic field perpendicular to the paper. $\omega_H$ is the angular gyro-frequency. It is supposed that the free orbit is three times as large as the forced orbit at all frequencies, although, in fact, the size of the forced orbit depends on frequency. (a) Extraordinary wave. It should be noticed that, when $\omega = \omega_H$, the forced orbit will become infinitely large and can no longer be less than the free one. (b) Ordinary wave.
Since \( \nu \) in both circles would be the same the combined orbit would have cusps, at which the combined velocity would be zero, and the starting point would be at one of these. The Extraordinary wave would produce a circular motion in the same sense for the 'forced' and the 'free' circles, and the Ordinary in opposite senses. A series of possible combined orbits is shown in fig. 14.3. With

\[
\omega = 4 \omega_H
\]

\[
\omega = \omega_H
\]

\[
\omega = \frac{1}{2} \omega_H
\]

(a) Extraordinary
(b) Ordinary

Fig. 14.3. Orbits described by an electron, originally at rest, when a rotating electric field is applied in the plane of the paper, and there is a steady magnetic field perpendicular to the paper. \( \omega \) is the angular frequency of the rotating field, and \( \omega_H \) is the angular gyro-frequency. The spiral orbit for the Extraordinary sense of rotation occurs when \( \omega \) and \( \omega_H \) are nearly, but not quite, equal, and it extends out to an infinite radius.
PART IV. SOME MISCELLANEOUS CONSIDERATIONS

CHAPTER 15

THE LORENTZ TERM

15.1. Introduction

It was at one time considered that an additional term, called the Lorentz Term, should be inserted in the equations of the magnetoionic theory, but the present view is that its inclusion is not correct. Although the theoretical reasoning, and the experimental evidence, on which it is generally excluded seem strong, it is nevertheless not overwhelming. Because there may still be some doubt on the point, and because many papers have been written in which the term is included, this chapter is devoted to a discussion of the subject.

15.2. History

In his book Lorentz (1909) considered the effect of passing an electro-magnetic wave through a medium which contained molecules, in each of which an electron could perform free damped oscillations. The equation of motion of one of these electrons under the action of the wave was of the form

\[ \dot{\mathbf{x}} + a \dot{\mathbf{x}} + b \mathbf{x} = e \mathbf{E}', \]  

(15.2.1)

where \( \mathbf{E}' \) was the electric field acting on the electron. Lorentz showed that \( \mathbf{E}' \) should not simply be equated to \( \mathbf{E} \), the field of the wave itself, but that it should also contain a contribution produced by the displacement of the electrons in the neighbouring molecules. He used an argument which was essentially as follows.

The medium which consisted of discrete molecules was assumed, for the purposes of calculation, to be replaced by a uniform continuous medium with properties which depended on the nature and number of the molecules. This was a ‘macroscopic’ approach of the type dealt with in Chapter 2. When the field produced by the medium at a single molecule was to be considered, the smoothing-out process of this macroscopic approach was reasonable at distances great compared with the inter-molecular distances, but was not necessarily reliable at distances smaller than that. Lorentz therefore suggested that the medium should be considered as a smoothed-out continuum everywhere outside a sphere of radius \( r \) surrounding the molecule under consideration, but inside this sphere the fields of the individual molecules should be considered in detail. The magnitude of \( r \) was to be such that it was great compared with the inter-molecular distances but small compared with the wave-length. The method of calculation is represented diagrammatically in fig. 15.1, where the microscopic theory is to be used inside the sphere and the macroscopic theory outside.

It was first shown that the field at the centre of the sphere, due to the external continuum alone, was independent of \( r \) and was given not by \( E' \) but by \( E + 4\pi P/3\varepsilon_0 \), where \( P \) was the volume polarisation of the continuum. Next the field due to the spherical distribution of discrete molecules was evaluated, and it was shown that, for the two special cases where the molecules were distributed \((a)\) at random and \((b)\) on a cubic lattice, the field at the centre was zero. For these two special cases, therefore, the effective field \( E' \) on the central molecule was that of the continuum alone and was given by

\[ E' = E + 4\pi P/3\varepsilon_0. \]  

(15.2.2)

Fig. 15.1. To illustrate the derivation of the Lorentz Term.
Magneto-Ionic Theory

The extra term \((4\pi P/\varepsilon_0)\) is known as the Lorentz Term. It is convenient to write it \(l(4\pi P/\varepsilon_0)\), so that when \(l = 0\) the term is not included and when \(l = \frac{1}{2}\) it is. When \(l = \frac{1}{2}\) Lorentz showed that the refractive index of a gas should be related to its density \(\rho\) by an equation of the form

\[
(\mu^2 - 1)/(\mu^2 + 2) \propto \rho. \tag{15.2.3}
\]

If \(l = 0\) there results a different relation, of a form which was earlier obtained by Sellmeier on an elastic-solid theory of the ether. This form of the theory leads to the expression

\[
(\mu^2 - 1) \propto \rho. \tag{15.2.4}
\]

Measurements of the refractive index of a gas at great pressures agree better with (15.2.3) than with (15.2.4). A theory in which \(l = 0\) is often called a Sellmeier theory, and one in which \(l = \frac{1}{2}\) is called a Lorentz theory.

When Appleton first developed the magneto-ionic theory he omitted the Lorentz Term, although he derived his expressions by a modification of those of Lorentz. The other early investigators derived the expressions without reference to Lorentz and also omitted the term. It was Hartree (1931a) who first suggested that Lorentz's argument might still be valid for the randomly distributed electrons in the ionosphere, and who drew attention to the way in which the equations would then be modified.

On simple arguments Hartree's suggestion seemed sound and the Lorentz Term was afterwards generally included in the expressions. In 1934, and later in 1943, Darwin examined the position critically for the case of free electrons embedded amongst positive ions, and he concluded that the inclusion of the term was not justified and it then became usual, once more, to omit it from the equations.

The only other theoretical suggestion is one made by Booker & Berkner (1938b), on more slender grounds, that \(l = 0\) when the amplitude of oscillation \(r\) of electrons is large compared with the mean distance \(a\) between them but that \(l = \frac{1}{2}\) when \(r < a\). When the electron density is \(10^6\text{cm}^{-3}\), \(a \approx 10^{-2}\text{cm.}\), and it is shown in Chapter 14 that a wave from a sender of 100 kW, can produce displacements of the electron of just about this magnitude. According to this view, therefore, the Sellmeier and the Lorentz theory might each be applicable under different practical circumstances.

Because a theoretical discussion of the problem is so involved, and has led, from time to time, to different conclusions, it has been thought desirable to devise experimental tests which might settle the point. These are discussed in §15.4. Before they are considered, however, it will be well to see how far the conclusions reached in this book on the basis of the Sellmeier theory \((l = 0)\) would need modification if the Lorentz Term were included \((l = \frac{1}{2})\).

15.3. Consequence of including the Lorentz Term \((l = \frac{1}{2})\)

When \(l = \frac{1}{2}\) it can be shown, by a straightforward modification of the calculations of Chapter 2, that the magneto-ionic equation (2.6.10) takes the form

\[
n^2 = 1 - \frac{X}{\left(1 + \frac{1}{2}X^2 - iZ\right) - \frac{1}{2}Y^2/(1 - \frac{3}{2}X - iZ)} \pm \left[\frac{1}{2}Y^2/(1 - \frac{3}{2}X - iZ)^2 + Y^2\right]^{1/2}. \tag{15.3.1}
\]

One result of making the change is that the values of \(X\) which make \(\mu = 0\), in the absence of collisions, are changed from \(X = (1 - Y)\), \(1\) and \((1 + Y)\) when \(l = 0\) to \(X = \frac{3}{2}(1 - Y)\), \(\frac{3}{2}\) and \(\frac{3}{2}(1 + Y)\) when \(l = \frac{1}{2}\). It follows that all electron densities, calculated with \(l = 0\) from observed values of the penetration frequencies, would have to be multiplied by \(\frac{3}{2}\) if \(l = \frac{1}{2}\). This change would not be important for most aspects of ionosphere theory.

The general shape of the \(\mu^2(X)\) curves, for different values of \(Y\), is much the same with \(l = 0\) or \(l = \frac{1}{2}\) provided \(\cos \theta < 1/\sqrt{3}\) (i.e. \(\theta > 55^\circ\)). If, however, \(\cos \theta > 1/\sqrt{3}\), there are some important differences when \(Y > 1\) and lies within certain prescribed ranges. These are indicated in fig. 15.2, in which the curves for the Sellmeier theory \((l = 0)\) are compared with those for the Lorentz theory \((l = \frac{1}{2})\). It is supposed in all these curves that \(\nu = 0\). Two values \(Y_1, Y_2\) of \(Y\) can then be found, such that, for \(1 < Y < Y_1\), there are two infinities in the \(\mu^2(X)\) curve for the lower sign (Extraordinary throughout since \(\nu = 0\)), for \(Y_1 < Y < Y_2\) the curves are similar to those for \(l = 0\); and for \(Y_2 < Y\) there are two infinities in the curve for the upper sign (Ordinary).

The magnitudes of \(Y_1\) and \(Y_2\) depend on the propagation angle \(\theta\).
Fig. 15.2. A comparison of the \( \mu(X) \) curves on the Sellmeier \((l = 0)\) and the Lorentz \((l = \frac{1}{2})\) theories. The magnitudes of \( Y_1 \) and \( Y_2 \) depend on the angle \((\theta)\) of propagation as shown in fig. 15.3.

They are given by the curves of fig. 15.3 which are derived from data given by Taylor (1933) and by Booker & Berkner (1938b). Although finite values of \( Y_3 \) are shown for \( \theta > 55^\circ \) they are of no practical significance because the corresponding infinite values of \( \mu \) are then reached for negative values of \( X \).

(a) **Oblique reflection from a layer**

Farmer & Ratcliffe (1936) pointed out that, ideally, an experimental test could be made in the following way. Suppose a wave of angular frequency \( \omega \) were incident obliquely, at an angle \( i \), on a layer of refractive index \( \mu \). Then it would just penetrate if

\[
\mu^2 = \sin^2 i. \quad (15.4.1)
\]

Now, in general, for the Ordinary wave component (\( QT \) approximation)

\[
\mu^2 = 1 - X/(1 + LX), \quad (15.4.2)
\]
where \( l = 0 \) on the Sellmeier theory and \( \frac{1}{4} \) on the Lorentz theory. If \( \omega_N \) is the angular plasma frequency appropriate to the electron density in the layer, a combination of (15.4.1) and (15.4.2) shows that the critical penetration frequency \( \omega_{ci} \) at the angle of incidence \( i \), is given by

\[
\frac{\omega_N^2}{\omega_{ci}^2} = \cos^2 i (1 - l \cos^2 i).
\]

(15.4.3)

Suppose now the critical penetration frequency is measured both at vertical incidence \( (i = 0) \) and at oblique incidence \( (i = i) \), and that the two frequencies \( \omega_0 \) and \( \omega_{ci} \) are compared, then (15.4.3) shows that

\[
\frac{\omega_{ci}}{\omega_0} = \sqrt{\left[ (1 - l)/(\sec^2 i - l) \right]} = \cos i \quad \text{when} \quad l = 0 \quad (15.4.4)
\]

\[
= \sqrt{\left[ 2/(3 \sec^2 i - 1) \right]} \quad \text{when} \quad l = \frac{1}{3}. \quad (15.4.5)
\]

The magnitudes of this quantity on the two theories are measurably different; thus if \( i = 60^\circ \) (15.4.4) and (15.4.5) give the following values

\[
\frac{\omega_{ci}}{\omega_0} = 0.5 \quad \text{Sellmeier}
\]

\[
= 0.42 \quad \text{Lorentz}.
\]

Unfortunately, this simple result cannot be used in practice because no ionospheric layer approximates sufficiently closely to the layer of uniform density assumed in the theory. If the electron density in the layer has some other height distribution then the maximum usable frequency (M.U.F.), for transmission over a given distance, is not necessarily the same as the penetration frequency. The M.U.F. could, ideally, be calculated for a transmission path of any given length, but the calculation would be involved because the direction of propagation would change along the path, and the full expression for \( \mu \) would have to be used in the calculations.

Attempts have been made to derive an approximate result by neglecting the effect of the earth's magnetic field, and using, for comparison, experimental results where the frequency was great in the hope that this neglect might not be too serious. Thus Ratcliffe (1939) calculated the M.U.F. to be expected for a layer with a parabolic distribution of electron density, and compared his results with the experimental measurements of Farmer, Childs & Cowie (1938) made on F-layer reflections over a distance of 464 km. The calculated difference between the two theories was so small (1%) that no decisive conclusion could be reached.

Smith (1941) developed a method of calculation in which a layer was considered to be made up of a number of superimposed uniform slabs. For typical \( E \) and \( F \) regions he found the results of fig. 15.4 in which, for a given value of \( \omega_0 \), the ratio of the M.U.F.'s on the two theories is shown as a function of distance. It is clear that, for the \( F \) layer, and distances of 500 km., the ratio is too small for a satisfactory test to be made, so that Ratcliffe's calculation might not have been expected to serve the purpose. Smith therefore chose to observe \( E \)-layer echoes, on 6 Mc./s., over a distance of 650 km. He observed the time of transition from \( E^- \) to \( F \)-layer reflection, as made evident by a change in the type of fading, and found the value of \( \omega_0 \) at this same time. From the results he concluded that, for the \( E \) layer, \( l = -0.02 \pm 0.05 \), a result which supports the Sellmeier theory.

Beynon (1947) extended the calculations of Ratcliffe, for a parabolic \( F \) layer, to greater distances, and showed that, for a transmission path of 1000 km., there is a difference of about 12% between the two theories in agreement with Smith's (1941) results, as shown in fig. 15.4. He also showed how to make an approximate correction for the curvature of the earth. He applied his results to the consideration of some experimental measurements of M.U.F. made on transmissions over distances of 1000 and 700 km.
The 1000 km. transmissions were from Germany during the war, and it was possible to measure \( \omega_\alpha \) only at the receiving end of the path. It was concluded that the values calculated from the Sellmeier theory differed from the measurements by about 1%, whereas those calculated from the Lorentz theory differed by 12%. With the shorter path, of 700 km., measurements of M.U.F., and of \( \omega_\alpha \) at both ends, were made simultaneously. Measured values differed from calculated (Sellmeier) values by \(-0.02\%\), whereas the Lorentz calculated values differed by 6%.

From these results it is usually supposed that the experimental evidence is in favour of the Sellmeier theory, and that the Lorentz Term should not be included. It must, however, be remembered that the effect of the earth's magnetic field has been omitted from the calculations, and it is quite possible that it might change the calculated results by an amount of the same order as the 12% or 6% on which the conclusion is based.

(b) *Phenomena near the gyro-frequency*

It has been shown in § 13.8 how an interesting echo-response, sometimes observed at frequencies less than the gyro-frequency, has been discussed by Booker & Berkner (1938 a, b) with reference to the Lorentz Term. They consider that the occurrence of this echo implies that the Lorentz Term should be included in the theory, but other workers have explained it in other ways. None of the explanations is entirely satisfactory.

(c) *Very low frequencies*

The explanation of propagation phenomena at very low frequencies depends on whether the Sellmeier or the Lorentz theory is used. The situation can be most easily understood by considering the forms taken by the curves (i) and (f) of fig. 15.2, appropriate to \( Y \gg 1 \), and modified so as to include the effects of collisions. When \( Y \gg 1 \) (e.g. when \( f = 16 \) kc./s. and \( Y = 80 \)), the electron density required to make \( X = 1 \) (for \( l = 0 \)), or \( X = \frac{3}{2} \) (for \( l = \frac{1}{2} \)), is so small (about 300 cm.\(^{-3}\)) that, at the appropriate level in the ionosphere, \( \nu > \omega_\alpha \) and the QL approximation is appropriate. If, above that level, it is supposed that \( X \) increases and \( Z \) decreases until it becomes unimportant, then the curves take the forms shown in fig. 15.5 (a) and (b). These curves are to be compared with those shown at (e) in fig. 13.3 and are to be explained in the same way.

The important difference between the two curves of fig. 15.5 is that in (a), for which \( l = 0 \), the refractive index of the Extraordinary wave increases steadily as \( X \) increases, whereas in (b), for which \( l = \frac{1}{2} \), the refractive index becomes infinite for a certain value of \( X \) (near \( X = 3 \)).

These results will first be applied to the phenomenon of 'whistlers' (Storey, 1953) which consist of waves of very low, audible, frequencies. A satisfactory theory suggests that they are propagated right through most of the ionosphere as Extraordinary waves, in a way which would be possible if the refractive index were given by the broken line in fig. 15.5 (a) \( (l = 0) \) but not if it were given by the broken curve in (b) \( (l = \frac{1}{2}) \), where no propagation is possible for values of \( X \) greater than that for which \( \mu = \infty \). The present theory of whistlers is thus satisfactory only if \( l = 0 \).

The second application is to very low-frequency waves (say 16 kc./s.) returned from the ionosphere by reflection at vertical incidence. Both theories, illustrated by curves (a) and (b) of fig. 15.5, suggest that the Ordinary wave, approximately circularly polarised with left-handed rotation, would be returned to the earth. The Sellmeier theory (a) suggests that the Extraordinary wave would not be returned, whereas the Lorentz theory (b) suggests (according to most writers) that it would be returned from the level where the refractive index is infinite. There is some doubt,
in fact, about whether reflection would be expected from this level (see §13.9), but if it were, the reflection of the Ordinary and Extra-
ordinary waves together would produce a resultant wave whose
polarisation would be complicated and variable. The experimental
fact that the returned wave, at very low frequencies, has a com-
paratively constant polarisation, nearly left-handed circular, has
led many workers to suggest that the Sellmeier theory alone can
explain the observations.

(d) Results from rocket experiments
Rockets have been used to measure simultaneously the Ordinary
and the Extraordinary refractive indices at a series of heights. From
the measured magnitude of the Ordinary index at each height the
magnitude of the Extraordinary index was deduced, and compared
with the observed magnitude. The agreement was good when the
Sellmeier theory was used, but was poor if the Lorentz theory
was used.

These results apply to waves of frequency about 4 Mc./s. travel-
ing in the E region.

16.1. Introduction
In the calculations of Chapter 2 it was assumed that if electrons
and massive ions were present in a medium in equal numbers then
only the electrons would influence the transmission of a wave,
because the magnitude of the quantity \(N/m\) for the electrons was
much greater than for the ions. In this chapter the question is
investigated in more detail, and expressions are derived to show
how, if the ions were sufficiently numerous, they would influence
the propagation of the wave.

The angular gyro-frequencies of the ions of mass \(m_i\) and electrons
of mass \(m_e\) are respectively equal to \(\mu_0 H_0 e/m_i\) and \(\mu_0 H_0 e/m_e\), so
that their ratio is \(m_e/m_i\). In the terrestrial ionosphere the gyro-
frequency of the ions is about 50 c./s., of the electrons about
1-25 Mc./s., and their ratio is about \(4 \times 10^{-5}\). The frequency of the
waves which will be considered here is always so much greater than
the gyro-frequency of the ions that the effect of the magnetic field
on the motion of the ions will be neglected. If the wave-frequency
were comparable with the gyro-frequency of the ions then the
calculations of this chapter would not be appropriate. Such cir-
cumstances might arise, for example, in the solar ionosphere near
a sunspot with a strong magnetic field.

16.2. The refractive index of a mixture
In this section the analysis of §§2.4 and 2.5 will be extended to
deal with a case where electrons and massive ions are present
together. The symbols will correspond to those in §§2.4 and 2.5,
and subscripts \(i\) and \(e\) will denote that they apply to ions and to
electrons, respectively.

The equations of motion of the electrons are given, as in the
absence of ions, by (2.4.3), (2.4.4) and (2.4.5) which are repeated
here:

\[
(e_0 X_e/4\pi) E_x = -P_{xe} + iY_e P_{xe} - iY_e P_{ie}, \tag{16.2.1}
\]
MAGNETO-IONIC THEORY

\[(\epsilon_0 X_d/4\pi)E_y = -P_{y\epsilon} + iY_LP_{x\epsilon}, \quad (16.2.2)\]
\[(\epsilon_0 X_d/4\pi)E_z = -P_{z\epsilon} - iY_TP_{x\epsilon}. \quad (16.2.3)\]

Because the action of the magnetic field on the ions can be neglected, the corresponding equations of motion for the ions are

\[(\epsilon_0 X_d/4\pi)E_x = -P_{x\epsilon}, \quad (16.2.4)\]
\[(\epsilon_0 X_d/4\pi)E_y = -P_{y\epsilon}, \quad (16.2.5)\]
\[(\epsilon_0 X_d/4\pi)E_z = -P_{z\epsilon}. \quad (16.2.6)\]

and, if there were no electrons present, the wave would have the refractive index \(n_t\) given by

\[n_t^2 = 1 + (4\pi/\epsilon_0)P_d/E. \quad (16.2.7)\]

This equation may be rewritten

\[P_d = (\epsilon_0/4\pi)(n_t^2 - 1)E. \quad (16.2.8)\]

The total polarisation \(\mathbf{P}\) is given by

\[\mathbf{P} = \mathbf{P}_t + \mathbf{P}_\epsilon, \quad (16.2.9)\]

and (2.3.3) requires that

\[P_z = P_{x\epsilon} + P_{y\epsilon} = -(\epsilon_0/4\pi)E_z. \quad (16.2.10)\]

Substitution from (16.2.8) then yields

\[(\epsilon_0/4\pi)E_z n_t^2 = -P_{y\epsilon}. \quad (16.2.11)\]

This is the first equation for which the motion of one type of particle is influenced by the presence of the other type, and it is of interest to notice how this 'coupling' between the two motions comes about. Under the joint action of the wave-field and the imposed magnetic field the electrons move, in general, in the \(Oz\)-direction, and their 'bunching' in this direction causes a field \(E_z\) which then acts upon the ions. This is the only force which the electrons exert upon the ions.* If it were not for this 'coupling' the two sets of particles would have made separate and independent contributions to the magnitude of the quantity \(n_t^2 - 1\), in the manner which is well known in the theory of the refractive index of mixed gases.

* If a Lorentz Term were included there would be another source of interaction. See Chapter 15.

WAVES IN MIXED MEDIUM

The calculations now proceed, as follows, in a manner parallel to that of § 2.5. Substitute from (16.2.11) into (16.2.3) to give

\[(1 - X_d/n_t^2)P_{x\epsilon} = -iY_TP_{x\epsilon}. \quad (16.2.12)\]

Now substitute for \(P_z\) in (16.2.1) to obtain

\[(\epsilon_0 X_d/4\pi)E_z = -\left(1 - \frac{Y_T^2}{1 - X_d/n_t^2}\right)P_{x\epsilon} + iY_LP_{y\epsilon}\]

or

\[(\epsilon_0 X_d/4\pi)(E_y/P_{y\epsilon}) = -\left(1 - \frac{Y_T^2}{1 - X_d/n_t^2}\right) - iY_LP_{y\epsilon}/P_{x\epsilon}, \quad (16.2.13)\]

and rearrange (16.2.2) to give

\[(\epsilon_0 X_d/4\pi)(E_y/P_{y\epsilon}) = -1 + iY_LP_{y\epsilon}/P_{x\epsilon}. \quad (16.2.14)\]

Now if a characteristic wave is to be propagated without change of form then, for reasons explained in Chapter 2, \(P_y/E_y = P_x/E_x\). But

\[P_x/E_x = (P_{x\epsilon} + P_{x\epsilon})/E_x = (\epsilon_0/4\pi)(n_t^2 - 1) + P_{x\epsilon}/E_x \quad (16.2.15)\]

from (16.2.8), and similarly

\[P_y/E_y = (\epsilon_0/4\pi)(n_t^2 - 1) + P_{y\epsilon}/E_y. \quad (16.2.16)\]

Equations (16.2.15) and (16.2.16) show that if \(P_x/E_x = P_y/E_y\) then also

\[P_{x\epsilon}/E_x = P_{y\epsilon}/E_y. \quad (16.2.17)\]

Under these conditions (16.2.13) and (16.2.14) lead to

\[iY_LP_{y\epsilon}/P_{x\epsilon}^3 - \{Y_T^2/2[1 - X_d/n_t^2]\}P_{x\epsilon}/P_{x\epsilon} + iY_L = 0. \quad (16.2.18)\]

Equation (16.2.18) is similar to (2.5.6) except that \(X\) is replaced by \(X_d/n_t^2\). Its solution gives the ratio \(r = P_{x\epsilon}/P_{y\epsilon}\), thus

\[r = -\frac{i}{Y_L}[\frac{1}{4}Y_T^2\frac{1}{1 - X_d/n_t^2} - \frac{1}{4}Y_T^2 + \frac{1}{4}Y_T^2]\], \quad (16.2.19)\]

which is similar to (2.5.8). Substitution into (16.2.14) now gives

\[(\epsilon_0 X_d/4\pi)(E_y/P_{y\epsilon}) = -1 + iY_L r \quad (16.2.20)\]

or

\[(4\pi/\epsilon_0)(P_{y\epsilon}/E_y) = -X_d/(1 - iY_L r). \quad (16.2.21)\]
Now from (2.3.12) the refractive index \( n \) is given in terms of the total polarisation by the expression

\[
n^2 = 1 + \left( \frac{4\pi}{e_0} \right) (P_y/E_y),
\]

where

\[
= 1 + \left( \frac{4\pi}{e_0} \right) (P_{yi} + P_{ye})/E_y
\]

(16.2.23)

from (16.2.7) and (16.2.21). Thus

\[
n^2/n_i^2 = 1 - (X_i/n_i^2)/(1 - iY_L r)
\]

(16.2.24)

\[
= 1 - \frac{X}{1 - \frac{1}{2} Y^2/(1 - X) \pm \left[ \frac{1}{4} Y^2/(1 - X) + Y^2 \right]^\frac{1}{2}}.
\]

(16.2.25)

If there were no massive ions \( n_i \) would be equal to unity, and (16.2.24) would be Appleton's equation for \( n^2 \) in terms of \( X \). Let us write this as \( A(X) \) so that

\[
A(X) = 1 - \frac{X}{1 - \frac{1}{2} Y^2/(1 - X) \pm \left[ \frac{1}{4} Y^2/(1 - X) + Y^2 \right]^\frac{1}{2}}.
\]

(16.2.26)

This form of the expression was given by Goubau (1935b).

The following calculation gives the polarisations of the waves:

\[
E_x/E_y = P_x/P_y = (P_{xi} + P_{xe})/(P_{yi} + P_{ye})
\]

(16.2.27)

\[
= r\left(1 + P_{xe}/P_{ye}\right)/(1 + P_{yi}/P_{ye}),
\]

(16.2.28)

where

\[
r = P_{xe}/P_{ye}.
\]

Now it has been shown (16.2.17) that

\[
E_x/P_{xe} = E_y/P_{ye} = B, \quad \text{(say)}
\]

and from (16.2.8)

\[
P_{xi} = (\epsilon_0/4\pi)(n_i^2 - 1)E_x
\]

\[
P_{yi} = (\epsilon_0/4\pi)(n_i^2 - 1)E_y
\]

so that

\[
P_{xi}/P_{xe} = (\epsilon_0/4\pi)(n_i^2 - 1)B
\]

and

\[
P_{yi}/P_{ye} = (\epsilon_0/4\pi)(n_i^2 - 1)B.
\]

Hence \( E_x/E_y = r \) and (16.2.27) becomes

\[
E_x/E_y = r.
\]

(16.2.28)

\( r \) thus represents the polarisation of the wave. It is given by (6.2.19) and can be related to the polarisation \( R(X) \) of a wave travelling in a magneto-ionic medium in which the ions are infinitely heavy. Equation (2.5.8) shows that

\[
R(X) = -i \left[ \frac{1}{2} Y^2 \right]/(1 - X \pm \left( \frac{1}{4} Y^2/(1 - X) + Y^2 \right)^\frac{1}{2}}.
\]

and comparison with (16.2.19) shows that

\[
E_x/E_y = r = R(X_i/n_i^2).
\]

16.3. The results for some special cases

It is now of interest to examine in detail the behaviour of the electron-ion mixture in some special cases. First it should be noticed that since the interaction between the two types of particles occurs only because of the longitudinal motion \( P_z \) of the electrons in the magnetic field, no interaction would be expected in those special cases where there is no motion of this kind. These special cases occur when the propagation is along the magnetic field, for both characteristic waves, and when it is perpendicular to the magnetic field for the Ordinary wave. These cases can be examined by the use of (16.2.24). If \( Y_L = 0 \) this yields

\[
n^2 = n_i^2 - X_d/(1 \pm Y_L)
\]

(16.3.1)

\[
= 1 - X_i - X_d/(1 \pm Y_L),
\]

(16.3.2)

and if \( Y_L = 0 \) it yields for the upper sign

\[
n^2 = n_i^2 - X_e
\]

(16.3.3)

\[
= 1 - X_i - X_e.
\]

(16.3.4)

Equations (16.3.1) and (16.3.2) show that the quantity \( n^2 - 1 \) can, in these cases, be derived by adding the magnitudes of \( n^2 - 1 \) for the electron medium and the ion medium separately. This is the type of addition which occurs when two dispersive gases are added.**

** In this statement it is assumed that the Lorentz Term is of negligible importance, as it usually is with gases at ordinary pressures. If this term were of importance in the ionosphere ions and electrons would not make independent contributions to the magnitude of \( n^2 \) - even in the special case considered here.
16.4. The effects of collisions in electron-ion mixtures

If it is supposed that the ions and the electrons collide with heavy particles with frequencies \( \nu_{i} \) and \( \nu_{e} \) respectively, then calculations similar to those of the preceding section lead to the results, first given by Goubau (1935b), that

\[
\frac{n_{e}^{2}}{n_{i}^{2}} = A(X_{e}/n_{i}^{2}),
\]

(16.4.1)

where

\[
A(X) = \frac{X}{1 - \frac{1}{iZ_{e} - \frac{1}{2} Y_{e}^{4}/(1 - X - iZ_{e})}} \pm \left[ \frac{1}{2} Y_{e}^{4}/(1 - X - iZ_{e})^{3} + Y_{e}^{4} \right]^{1/2},
\]

(16.4.2)

and

\[
\frac{n_{e}^{2}}{n_{i}^{2}} = 1 - X_{i}/(1 - iZ_{i}),
\]

\[
Z_{e} = \nu_{e}/\omega,
\]

\[
Z_{i} = \nu_{i}/\omega.
\]

As before the complex refractive index \( n \) is given in terms of the function \( A(X) \) which gives its magnitude for electrons alone. \( n_{i} \) is the complex refractive index which the medium would have if the ions were alone present. It may also be shown that the polarisation is given by

\[
r = E_{x}/E_{y} = -\frac{i}{Y_{e}} \left[ \frac{1}{2} Y_{e}^{3}/(1 - X_{i}/n_{i}^{2} - iZ_{e})^{3} \left( \frac{1}{2} Y_{e}^{4}/(1 - X_{i}/n_{i}^{2} - iZ_{e})^{3} + Y_{e}^{4} \right)^{1/2} \right],
\]

(16.4.3)

where \( R(X_{i}/n_{i}^{2}) \) is the function (2.6.7) which describes the polarisation for a pure electron gas.
CHAPTER 17

INHOMOGENEOUS MEDIA

17.1. Introduction

Most of this book has been concerned with the propagation of waves through a magneto-ionic medium which is homogeneous. When there is a gradient of electron density or of collision frequency or of both, as there is in the ionosphere, many interesting phenomena can occur. Their complete treatment requires a detailed solution of the appropriate differential equations and is quite complicated. In this chapter a descriptive account will be given of some of the results and some simple physical explanations will be given.

It will be supposed that, in the medium considered, the imposed magnetic field is constant, that the electron density and collision frequency are stratified parallel to the x-y-plane, and that they have a gradient in the z-direction, which is taken to be vertical.

17.2. Conditions for reflection at vertical incidence (no imposed field)

First consider a wave travelling vertically upwards, along positive Oz, and suppose that the electron density increases steadily upwards. If there are no collisions between electrons and heavy particles the refractive index \( \mu \) will become zero at some level at which, according to a ray theory, the wave will be completely reflected. If the situation is investigated more fully, by considering the appropriate differential equation, it is again found that the wave is completely reflected as if from the level where \( \mu = 0 \) but with a phase-shift which depends on the gradient. This phase-shift is not expected on a ray theory. Above the level where \( \mu = 0 \) there is an evanescent wave, but no propagation of energy.

Next consider a 'layer' of electrons, over-dense for the frequency concerned, so that there are two levels, as at \( AA' \) and \( BB' \) in fig. 17.1, between which the complex refractive index \( (n) \) is purely imaginary (i.e. \( n^2 \) is negative, \( \mu = 0 \), and \( \gamma \) is finite). Then if the distance \( AB \) is not too great the evanescent wave may not die away completely before the upper level \( BB' \) is reached, and some energy may be transmitted. The phenomenon is inexplicable on ray theory; it corresponds to the 'tunnel effect' in wave mechanics.

The fractions of the energy transmitted or reflected depend on the distance \( AB \) and on the distribution of the electrons throughout this distance. When back and forth reflection of the waves between the levels \( AA' \) and \( BB' \) is considered it is found that the transmission and reflection coefficients are usually periodic functions of the distance \( AB \). Detailed calculations have been made for layers having different specified electron distributions; they are useful in considering reflection from thin layers, such as might possibly be responsible for \( E_r \) reflections.

If it is next supposed that the electrons collide with heavy particles there are two changes. First, the refractive index never falls to zero, and second, absorption occurs.

Suppose that, over a limited portion of a medium where the mean refractive index is \( \mu \), there is a linear gradient of refractive index, and approximate, as in fig. 17.2 (a), by supposing that it is made up of a series of equal infinitesimal steps, each of thickness \( \delta x \), in each of which the refractive index changes by \( \delta \mu \). Suppose also that a wave of unit amplitude is travelling vertically upwards. Then the simplest wave-theory, of the type used in deriving Fresnel's reflection coefficients, shows that each boundary reflects a weak wavelet with amplitude \( \delta \mu / 2 \mu \). Suppose that these wavelets are so weak
that the original wave is effectively undiminished as it travels and
that multiple reflections of the wavelets are negligible. Then,
as indicated in fig. 17.2(b), the reflected wave is the resultant
of all the partially reflected wavelets, when they have been added
with due regard to their phase. Now if \( \lambda_0 \) is the free-space wave-length
the reflected wavelet is retarded in phase by \((2\pi/\lambda_0)(2\mu\delta z)\) at each
successive step. The addition of these wavelets is represented on
an amplitude-phase diagram by a regular polygon with sides of
length \( \delta\mu/2\mu \) having an angle \((2\pi/\lambda_0)(2\mu\delta z)\) between them. In
the limit, when \( \delta z \) and \( \delta\mu \) tend to zero, the polygon becomes a
circle in which the relation between an element \( (da) \) of length and
the element \( (d\phi) \) of angle subtended at the centre is given by

\[
da/d\phi = L_t \{ \delta\mu/2\mu \} + \{(2\pi/\lambda_0)(2\mu\delta z)\} \\
= (1/8\pi)(\lambda_0/\mu^2)(d\mu/dz). 
\]  

(17.2.1)

If there is some absorption the main wave will weaken as it travels
so that successive wavelets are weaker, and the circle will become
a spiral as in fig. 17.2(c). The resultant reflected wave will then be
represented by the radius \( OR \) which is equal to \((1/\pi)\) times the
semi-circumference \( OST \). But \( OST \) is given by

\[
OST = \int_0^\mu (da/d\phi) d\phi \\
= (\lambda_0/8\mu^2)(d\mu/dz) 
\]

(17.2.2)

from (17.2.1). The resultant reflected wave is hence given by

\[
OR = (\lambda_0/8\pi\mu^2)(d\mu/dz). 
\]

(17.2.3)

Equation (17.2.3), derived by admittedly crude reasoning, gives
a correct account of what determines the magnitude of the wave
reflected from a gradient of refractive index (see, for example,
Gans (1915)).

There must be several convolutions of the spiral for (17.2.3) to
hold, i.e. the gradient must extend over several wave-lengths as
measured in the medium. If this restriction is removed results
which are correct in order of magnitude can be deduced by con-
sidering a portion of one turn. It then becomes apparent, for
example, that a linear gradient, extending over a limited distance,
produces a reflected wave of amplitude which depends jointly on
the gradient and on the distance measured in wave-lengths-in-the-
medium.

If the 'gradient reflection coefficient' given by (17.2.3) is very
small compared with unity only the incident wave need be con-
sidered, and 'ray optics' will suffice for the solution of the problem.*
This statement can be more rigidly substantiated by a detailed
consideration of the appropriate differential equation (see refer-
ences in the Bibliography).

Equation (17.2.3) shows that a simple ray treatment is most
likely to be misleading:

(a) when the free-space wave-length \( \lambda_0 \) is great, i.e. for long
waves,

(b) near the level where \( \mu \) approaches zero, and

* It is interesting to notice that, in wave mechanics, the corresponding
conditions determine whether or not Newtonian mechanics is sufficient for the
solution of a problem.
(c) when $d\mu/dz$ is large. This can be written $(d\mu/dN)(dN/dz)$ and shows that ray theory may be inadequate either because $(d\mu/dN)$ is great or because there is a large gradient of $N$.

In Chapters 12 and 13 it was shown that, in the terrestrial ionosphere, $\mu$ is small and $(d\mu/dN)$ is large near the level where $\mu$ would be zero if there were no collisions. The simple treatment given here suggests, correctly, that strong reflection would be expected near this level.

### 17.3. Coupling between characteristic waves

If a characteristic wave is launched in a homogeneous magneto-ionic medium it will keep the same polarisation and will remain characteristic. If there is a gradient of electron density, collision frequency, or magnetic field, the polarisation appropriate to a characteristic wave will change from place to place, as shown, for example, in fig. 7.5, and if the wave is to remain characteristic its polarisation must change also. If the medium changes too rapidly this cannot always happen, and what was originally one characteristic wave will become a sum of two. Under these conditions the second characteristic wave is said to be produced from the first by a process of coupling. A full treatment of the phenomenon has been given in papers which develop the matter in terms of coupled differential equations. The essentials of the matter can be understood by simple physical arguments, as follows.

Suppose that a single characteristic wave (say the Ordinary) is incident normally on a boundary which separates two homogeneous magneto-ionic media which have different electron densities. The complex polarisation ($R_0$) of the Ordinary wave will, in general, be different on the two sides of the boundary, so that, in order to satisfy the boundary conditions, it will be necessary to include some proportion of Extraordinary wave, on both the transmission and the reflection sides. It will be said that both these Extraordinary waves have been produced by a process of ‘coupling’ at the boundary.

The waves produced by the coupling will be comparatively strong under either of the following conditions:

(a) if the difference $\Delta R_0$ between the complex Ordinary polarisations in the two media is great, or

(b) if the difference $(R_0 - R_x)$ between the complex polarisations of the Ordinary and the Extraordinary waves is small, since then, in order to correct the failure of the boundary conditions when Ordinary waves are alone used, it will be necessary to add Extraordinary waves with large amplitudes.

Suppose next that the wave is passing through a medium which consists of a series of thin slabs, in each of which the electron density is slightly different, and consider what would happen in the limit when there was a continuous gradient of density. The sum of the Extraordinary waves produced at each interface would, when due regard had been paid to phase, constitute the large-scale transmitted and reflected waves produced by the coupling. They would, in general, be so small that it would be unnecessary to consider second-order waves which would arise by coupling from these waves themselves.

In the backward direction the total Extraordinary wave has an amplitude which can be estimated by an argument similar to that of the last section, and which depends chiefly on the gradient $(dR_0/dz)$ of the Ordinary polarisation and the difference $(R_0 - R_x)$ of the two characteristic polarisations. Full calculations (e.g. Försterling, 1942) show that the amplitude of the reflected wave is proportional to a coupling parameter $\psi = (dR_0/dz)/(R_0 - R_x)$, which may be written

$$\psi = (dR_0/dz)/(R_0^2 - 1),$$

(17.3.1) since $R_0 R_x = 1$.

In the forward direction the amplitudes of the Extraordinary wavelets produced at the discontinuities are determined by the same considerations, but the question of their phases is more important. Suppose the Ordinary wave travels along positive $Ox$ and at a plane $z_1$, where its phase is proportional to $\mu_0 z_1$, suppose it produces an Extraordinary wave by a process of coupling. Then on arrival at a further plane at a distance $z$ ($> z_1$) the Extraordinary wave will have a phase proportional to

$$\mu_0 z_1 + (z - z_1) \mu_x = z \mu_x + z_1 (\mu_0 - \mu_x).$$

The resultant Extraordinary wave at the plane $z$ is the sum of all the wavelets produced by coupling at the planes at the different distances $z_1$, and if $(\mu_0 - \mu_x)$ is small their phases are nearly the
same and the resultant wave is large. The transmitted wave originating in coupling can thus be large either because $\psi$ is large or because $\mu_0 - \mu_2$ is small.

The backward- and forward-travelling waves produced by coupling may be important in the ionosphere in ways which are described in the next two sections.

17.4. The coupling echo

A relatively strong reflected wave, of the complementary characteristic polarisation, would be expected from a level where $\psi$ (17.3.1) has a large value. This is most likely to occur when $R_0 = R_x$, i.e. where $R_0 = R_x = 1$, and reference to § 7.2 shows that the appropriate conditions are that $X = 1$ and $\nu = \omega_0$. An echo, originating at this level through a process of coupling, has been called a 'coupling echo'.

With the model ionosphere represented in fig. 13.13 $|\omega_0/\nu| = 1$ at the level 93 km, and $X = 1$ at this level for an angular frequency of about $10^6$ sec$^{-1}$, i.e. a wave frequency of 160 kc/s. The workers at Pennsylvania State University believe that they have observed echoes, of the type discussed here, on a frequency of 150 kc/s. They appear somewhat infrequently from a level of 93 km, and their suggestion is that when they appear the ionospheric structure is such that $X = 1$ at the level where $\nu = |\omega_0|$.

17.5. Emergent polarisation

When a plane-polarised wave is sent into the ionosphere it is split into the two characteristic waves. Under certain conditions only one of these is returned to the ground, for example, on high frequencies of a few megacycles per second the maximum electron density may be insufficient to return the other, or on a lower frequency the refractive index for the other wave may never fall to a low enough value to produce reflection (see figs. 13.14 (d) and (e)). Now if the wave which returns to the ground continually maintained the polarisation appropriate to the level where it finds itself it would emerge with a polarisation appropriate to the condition $N = 0$, $\nu = \nu_0$, where $\nu_0$ is the value of $\nu$ at the level where $N = 0$.

There are, however, two different reasons why this polarisation is not always the one observed.

(a) Passage through a region of strong coupling

It was shown in § 7.2 that the polarisation ellipses of both characteristic waves must lie with their major axes in the same pair of quadrants, which, in the Northern Hemisphere, is the N.E.-S.W. pair, and this restriction would apply to the polarisation of an emergent wave if it consisted of a single characteristic wave. If, however, on its way to the ground, a single characteristic wave passes through a region of strong coupling, where the quantity $(dR/d\zeta)/(R^2 - 1)$ is comparatively large, the other characteristic wave is excited in appreciable magnitude. The two waves may then travel onward to the ground, each changing its polarisation as it travels so as to remain characteristic for the conditions where it finds itself, and each travelling with its own phase velocity. On emerging from the ionosphere the two waves will combine to produce a resulting polarisation which is determined by the ratio of the two amplitudes, by the nature of the two characteristic polarisations, and by the relative phases of the two waves. The resulting polarisation ellipse may have its major axis outside the N.E.-S.W. pair of quadrants in a direction 'forbidden' by the simple magneto-ionic theory.

The workers at Pennsylvania State University have found, on a frequency of 150 kc/s., that the direction of the polarisation ellipse of the emerging wave sometimes changes rapidly. They have called this phenomenon a 'flip-over' and believe that it occurs when the level where $\nu = \omega_0$ is also the level where $X = 1$ for the appropriate frequency, so that the coupling is relatively great. (See § 13.10.)

(b) Limiting polarisation

If the curves of fig. 7.5 are examined it will be seen that the polarisations of the characteristic waves for the conditions $X = 0$, $\nu = \nu_0$ are determinate and are given by the value of $R$ when $\xi = -\omega/|\omega_0|$. This implies that a medium which does not in the limit contain any electrons ($X = 0$) can transmit only these two
polarisations without change of form. But we know that, in fact, such a medium will transmit any polarisation unaltered. How does this come about?

The answer is that the medium can, indeed, be thought of as splitting the incident wave up into the two characteristic waves, each transmitted with its own phase velocity, but when \( N = \infty \) the two phase velocities are the same, so that, at all distances, their combination results in a wave which has the same polarisation as the original one.

The following question now arises. Suppose a characteristic wave, nearing the end of its journey out of the ionosphere, passes through a final part of its path where \( N \) is very small but finite and where \( \nu \) changes from \( \nu_1 \) to \( \nu_0 \). If an attempt were made to decide what would happen to the polarisation by referring to fig. 7.5 a fixed value of \( \xi \) would be taken to represent \((\omega/|\omega_i|)(X-1)\) (it would be nearly equal to \((\omega/|\omega_i|))\), and the change of \( R \) would be examined corresponding to the change of the parameter \((|\omega_i|/\nu)\) (marked on the curves) as \( \nu \) varied from \( \nu_1 \) to \( \nu_0 \). This would lead to the conclusion that the emerging polarisation would be that appropriate to \( \nu_0 \). But it is clear, from what was said in the previous paragraph, that, because \( N \rightarrow \infty \), the two refractive indices are nearly the same, and the resulting polarisation cannot change as the wave travels, so that, in fact, the emergent polarisation would correspond to \( \nu_1 \).

Arguments along these lines suggest that the 'limiting polarisation' is determined, not by the value \( \nu_0 \) of \( \nu \) when \( N \rightarrow \infty \), but by its value at some greater height, which depends on the difference of the refractive indices.

It was shown in § 17.3 that the transmitted wave produced by coupling can be large either because \( \psi \) (17.3.1) is large, or because \((|\mu_0 - \mu_2|)\) is small. Usually when a characteristic wave (for example, the Ordinary) traverses the lower part of the ionosphere \( \psi \) is small, but, at a certain level where the electron density and collision frequencies have the values \( N_1 \) and \( \nu_1 \), the quantity \((|\mu_0 - \mu_2|)\) becomes small and an appreciable wave of complementary polarisation (Extraordinary) arises through the coupling process. As the wave travels out of the ionosphere \( N \) and \( \nu \) change and the polarisation of the Ordinary wave changes as it travels, so as to remain characteristic for the place where it is, but, at each level, the added Extraordinary is just sufficient, in amplitude and phase, to combine with the Ordinary to maintain the polarisation appropriate to the level \((N = N_1 \text{ and } \nu = \nu_1)\) where this process started.

It is difficult to decide precisely what values \( N_1 \) and \( \nu_1 \) determine the emerging polarisation, but some first attempts, listed in the Bibliography, have sufficed to show that it is not usually sufficient to suppose that the emerging polarisation is correctly determined by taking the magnitude of \( \nu \) at the level where \( N \) becomes very small.
CHAPTER 18

A MAGNETO-IONIC MEDIUM CONSIDERED AS AN ANISOTROPIC DIELECTRIC, AND A COMPARISON WITH CRYSTAL OPTICS

18.1. Introduction

A magneto-ionic medium is anisotropic in the sense that, in general, an electric field applied in one direction produces a resulting polarisation in a different direction. Crystals provide another example of anisotropic media, and the travel of electro-magnetic waves through them is fully discussed in text-books on optics. In this section the propagation of waves through a magneto-ionic medium and through a crystal are compared and contrasted, and it is shown that, although there are many similarities, propagation through a magneto-ionic medium is more complicated than propagation through a crystal.

18.2. Calculation of refractive index using tensor notation

The anisotropic nature of the magneto-ionic medium is represented by the constitutive relations of (2.6.5). It is convenient to make these more general by supposing that the imposed field \( \mathbf{H}_0 \) lies in a general direction, and writing \( Y_x, Y_y \) and \( Y_z \) to represent the components of the vector \( \mathbf{Y} = (\mu_0 e/4\pi) \mathbf{H}_0 \) along the coordinate axes. If, also, we write \( U \equiv 1 - iZ \), then the equations (2.6.5) are replaced by

\[
\left( \frac{\varepsilon_0 X}{4\pi} \right) E_x = -UP_x - iY_xP_y + iY_yP_z, \\
\left( \frac{\varepsilon_0 X}{4\pi} \right) E_y = iY_xP_z - UP_y - iY_zP_x, \\
\left( \frac{\varepsilon_0 X}{4\pi} \right) E_z = -iY_xP_z + iY_yP_x - UP_x. \\
\]

(18.2.1)

These equations may be written in matrix form

\[
\begin{pmatrix}
E_x \\
E_y \\
E_z
\end{pmatrix} = \frac{4\pi}{\varepsilon_0 X} \begin{pmatrix}
-U & -iY_x & iY_y \\
iY_x & -U & -iY_z \\
-iY_y & iY_z & -U
\end{pmatrix} \begin{pmatrix}
P_x \\
P_y \\
P_z
\end{pmatrix}. \\
\]

(18.2.2)

where

\[
\sigma^{-1} = \frac{4\pi}{\varepsilon_0 X} \begin{pmatrix}
-U & -iY_x & iY_y \\
iY_x & -U & -iY_z \\
-iY_y & iY_z & -U
\end{pmatrix}. \\
\]

(18.2.6)

is the 'inverse susceptibility matrix'.

To determine the refractive index we next make use of Maxwell's equations which, for a wave travelling along \( Oz \), can be put in the form of (2.3.11), (2.3.12) and (2.3.3), which are

\[
(e_0/4\pi)(n^2 - 1) E_x = P_x, \\
(e_0/4\pi)(n^2 - 1) E_y = P_y, \\
-(e_0/4\pi) E_z = P_z. \\
\]

(18.2.7)

(18.2.8)

(18.2.9)

or in matrix form

\[
\frac{e_0}{4\pi} \begin{pmatrix}
n^2 - 1 & o & o \\
o & n^2 - 1 & o \\
o & o & -1
\end{pmatrix} \begin{pmatrix}
P_x \\
P_y \\
P_z
\end{pmatrix} = \begin{pmatrix}
P_x \\
P_y \\
P_z
\end{pmatrix}. \\
\]

(18.2.10)

For an isotropic medium substitution for \( E_x \) from (18.2.4) into (18.2.7) gives

\[
(e_0/4\pi)(n^2 - 1) \sigma^{-1} = I. \\
\]

(18.2.11)

In a similar way for the anisotropic medium substitution of \( E \) from (18.2.5) into (18.2.10) gives

\[
\frac{e_0}{4\pi} \begin{pmatrix}
n^2 - 1 & o & o \\
o & n^2 - 1 & o \\
o & o & -1
\end{pmatrix} \sigma^{-1} = I, \\
\]

(18.2.12)

or, after substituting from (18.2.6),

\[
\frac{1}{X} \begin{pmatrix}
n^2 - 1 & o & o \\
o & n^2 - 1 & o \\
o & o & -1
\end{pmatrix} \begin{pmatrix}
-iY_x & iY_y \\
iY_x & -U & -iY_z \\
-iY_y & iY_z & -U
\end{pmatrix} = I. \\
\]

(18.2.13)
This implies
\[
\begin{align*}
-U(n^2 - 1) - X &\quad -iY_z(n^2 - 1) &\quad iY_z(n^2 - 1) \\
iY_z(n^2 - 1) &\quad -U(n^2 - 1) - X &\quad -iY_z(n^2 - 1) \\
iY_z &\quad -iY_z &\quad U - X
\end{align*}
\]
\begin{equation}
= 0, \quad (18.2.14)
\end{equation}

or, if we write \(A \equiv X/(n^2 - 1)\),
\[
\begin{align*}
-U - A &\quad -iY_z &\quad iY_z \\
iY_z &\quad -U - A &\quad -iY_z \\
iY_z &\quad -iY_z &\quad U - X
\end{align*}
\begin{equation}
= 0. \quad (18.2.15)
\end{equation}

If now we write
\[
Y_y \equiv 0, \quad Y_z \equiv Y_{L}, \quad Y_x \equiv Y_r,
\]
this gives
\[
A^2 + 2A(U - \frac{1}{2}Y^2/(U - X)) + U(U - Y^2/(U - X)) - Y^2 = 0.
\begin{equation}
(18.2.16)
\end{equation}

This equation is a quadratic in \(n^2\) from which the two values of \(n^2\) appropriate to the two characteristic waves can be deduced. It is the same as Appleton's equation (2.6.10).

18.3. The dielectric and susceptibility tensors

It is interesting to consider the nature of the tensor susceptibility \(\sigma\), derivable by inverting \(\sigma^{-1}\) (18.2.6) to give
\[
\sigma = \frac{-\varepsilon_0 X}{4\pi U(U^2 - Y^2)} \left( \begin{array}{ccc}
U^2 - Y^2 & iUY_z - Y_i^2 & iUY_z - Y_i^2 \\
iUY_z - Y_i^2 & U^2 - Y^2 & iUY_z - Y_i^2 \\
iUY_z - Y_i^2 & iUY_z - Y_i^2 & U^2 - Y^2
\end{array} \right).
\begin{equation}
(18.3.1)
\end{equation}

If \(H_0\) is taken to lie along the Ox-direction, so that \(Y_x = Y_y = 0\), this gives
\[
\sigma = \frac{\varepsilon_0}{4\pi} \left( \begin{array}{ccc}
CU & -iCY_z & 0 \\
iCY_z & CU & 0 \\
0 & 0 & -X/U
\end{array} \right),
\begin{equation}
(18.3.2)
\end{equation}

where \(C \equiv X/(U^2 - Y^2)\). If now we write the dielectric tensor \(\epsilon\) we have
\[
\epsilon = 1 + (4\pi/\varepsilon_0) \sigma,
\begin{equation}
(18.3.3)
\end{equation}

and (18.3.2) can be rewritten
\[
\epsilon = \left( \begin{array}{ccc}
1 + CU & -iCY_z & 0 \\
iCY_z & 1 + CU & 0 \\
0 & 0 & 1 - X/U
\end{array} \right).
\begin{equation}
(18.3.4)
\end{equation}

The dielectric tensor of (18.3.4) can be expressed simply in terms of the quantities \(n^2_1, n^2_2, n^2_3\), which are respectively the squares of the refractive indices for the Ordinary and Extraordinary waves propagated along the field, and the Ordinary wave propagated across the field, as given by
\[
\begin{align*}
n^2_1 &= 1 - X/(U + Y), \\
n^2_2 &= 1 - X/(U - Y), \\
n^2_3 &= 1 - X/U.
\end{align*}
\begin{equation}
(18.3.5)
\end{equation}

From (18.3.5)
\[
\frac{1}{2}(n^2_1 + n^2_2) = 1 + CU, \\
\frac{1}{2}(n^2_1 + n^2_3) = -CY,
\begin{equation}
(18.3.6)
\end{equation}

so that (18.3.4) becomes
\[
\epsilon = \left( \begin{array}{ccc}
\frac{1}{2}(n^2_1 + n^2_2) & \frac{i}{2}(n^2_1 - n^2_2) & 0 \\
\frac{i}{2}(n^2_1 - n^2_2) & \frac{1}{2}(n^2_1 + n^2_3) & 0 \\
0 & 0 & n^2_3
\end{array} \right).
\begin{equation}
(18.3.7)
\end{equation}

This equation gives the dielectric tensor \(\epsilon\) in terms of the three refractive indices \(n^2_1, n^2_2, n^2_3\) for the three simple characteristic waves.

18.4. The principal directions

Anisotropic media with dielectric tensors are familiar in the theory of crystal optics, but there the tensors are of a simpler form and can be written
\[
\epsilon = \left( \begin{array}{ccc}
n^2_1 & 0 & 0 \\
0 & n^2_2 & 0 \\
0 & 0 & n^2_3
\end{array} \right),
\begin{equation}
(18.4.1)
\end{equation}

where \(n_1, n_2, n_3\) are the refractive indices of waves travelling along the three principal directions of the crystal. It is interesting to inquire whether there are three principal directions in a magneto-ionic medium such that the dielectric tensor takes the form (18.4.1) when referred to them as axes. It can be shown that (18.3.7) can be transformed into (18.4.1) if it is rewritten with reference to three principal directions, two of which are complex, and the third is the direction of the imposed magnetic field. The quantities \(n_1, n_2, n_3\) are, as before, the refractive indices of the 'principal' waves given by (18.3.5).
18.5. The wave-velocity surfaces and the $\mu(\theta)$ surfaces

When a medium, such as a crystal, is described by a dielectric tensor like (18.4.1), two characteristic waves can, in general, travel along any given direction, and it is usual to represent their properties by constructing a pair of wave-velocity surfaces such that the radii vectors are proportional to the wave velocities, i.e. inversely proportional to the refractive indices. For an ordinary crystal these take the form shown in fig. 18.1, where $Ox$, $Oy$, $Oz$ are the principal directions.

Fig. 18.1. The wave-velocity surfaces for a bi-axial crystal. These are surfaces of $1/\mu$. $OA$ and $OA'$ are the two optic axes, the inner and outer surfaces are in contact along these lines.

Clemmow & Mullaly (1955) have investigated the form of the wave-velocity surfaces when the electrons do not make collisions ($v = 0$), and have shown that they are more complicated than those in crystals. Since it usual in the magneto-ionic theory to consider the refractive index rather than the wave-velocity they calculated the surfaces which gave $\mu(\theta)$ in different directions $\theta$, rather than those which gave the wave-velocity ($1/\mu$). These are surfaces of revolution about the imposed magnetic field $H_0$, and their sections by planes containing $H_0$ can take a series of different shapes which are shown in their paper. A few examples are shown in figs. 18.2 and 18.3 for the purpose of illustration. They have been chosen to correspond to conditions already illustrated in figs. 6.8(b) and 6.9(b) in which $\mu(X)$ was plotted when $Y = 0.5$ and $Y = 2$. In figs. 18.2 and 18.3 these curves are replotted in terms of $\mu(X)$ instead of $\mu^*(X)$. As in figs. 6.8 and 6.9 the curves which represent longitudinal and transverse propagation are labelled $L$ and $T$ respectively, and curves for intermediate directions of propagation lie in one or other of the cross-hatched regions, for the upper and lower signs in the magneto-ionic equation.

Fig. 18.2. To illustrate some $\mu(\theta)$ surfaces for the special case where $Y = 0.5$. (a) shows the regions (shaded) in the $\mu$, $X$-plane within which the curves must lie for any value of $\theta$. These curves should be compared with those of fig. 6.8(b) which represent the same facts in the $\mu^*$, $X$-plane. (b) shows the $\mu(\theta)$ surfaces when $X = 0.45$. (c) shows the $\mu(\theta)$ surfaces when $X = 0.95$. The letters in (b) and (c) correspond to those in (a). $H_0$ indicates the direction of the imposed magnetic field.
Fig. 18.2 (b) shows the $\mu(\theta)$ surfaces when $Y = 0.5$ and $X = 0.45$, corresponding to the line CDEF in fig. 18.2 (a). As $\theta$ varies from 0 to $\frac{\pi}{2}$, $\mu$ for the extraordinary wave changes from the value marked C on both figures to that marked D and $\mu$ for the ordinary wave changes from F to E.

In the same way fig. 18.2 (c) shows the $\mu(\theta)$ surfaces when $Y = 0.5$ and $X = 0.95$, corresponding to the line GKMN in fig. 18.2 (a). As before, corresponding points on the two figures are marked with corresponding letters. Here there is a real value of $\mu$ for the lower sign only if the direction of propagation does not approach too closely that of the imposed magnetic field.

Fig. 18.3 shows corresponding curves when $Y = 2$. The $\mu(X)$ curves of fig. 18.3 (a) correspond to those of fig. 6.9 (b). This time it will be seen that, when $X = 1.2$, $\mu$ for the upper sign is real only when the direction of propagation is sufficiently nearly along the imposed field.

The $\mu(\theta)$ surfaces of figs. 18.2 and 18.3 have been chosen, from amongst those listed by Clemmow & Mullaly, for purposes of illustration only. Other types of surface can occur, as explained in their paper.

### 18.6. The optic axes

The wave-velocity surfaces of fig. 18.1, appropriate to a wave travelling through a crystal, are in contact along two directions $OA$ and $OA'$, so that when the wave-normal is along these directions the two characteristic waves have the same velocity, and the directions $OA$ and $OA'$ are called the optic axes. In a magneto-ionic medium in which the electrons do not make collisions the wave-velocity surfaces, or the $\mu(\theta)$ surfaces, nowhere touch, so that there are no optic axes. Examples of the way in which the two surfaces are quite distinct are shown in figs. 18.2 and 18.3.

When collisions are taken into account, however, the refractive indices are altered, and for some special conditions the two $\mu(\theta)$ surfaces touch, and the two waves have the same wave-velocity. This occurs in a medium and for a frequency which makes $X = 1$, and in a direction $\theta$ such that $\frac{1}{2} \sin^2 \theta/\cos \theta = v/\omega_H$ or $\nu = \omega_p$. This is, of course, the important condition, considered in detail in Chapter 7, which makes the polarisations of the two waves the same. Along the optic axis direction there is thus no distinction between the two characteristic waves. In the terminology of § 17.4 the two waves are strongly coupled and, indeed, the condition $(X = 1, \nu = \omega_p)$ of critical coupling is precisely the condition that there is an optic axis and that the waves are travelling along it.

The form of the wave surfaces when collisions occur and when there is an optic axis direction, can be illustrated from the curves of
fig. 10.2. The line in this figure corresponding to \( Z = 0.01 \) shows how the refractive indices \( \mu \) change with angle. The condition \( \nu = \omega_0 \) is the same as \( Z = \omega_0 / \omega \) and occurs when \( \theta = 8^\circ \). If the values of \( \mu \) shown in fig. 10.2 when \( X = 1 \) and \( Z = 0.01 \) are plotted in a polar curve, it takes the form shown roughly in fig. 18.4. This figure emphasises the fact that, for a magneto-ionic medium, there are not just two optic axes, but there is a cone of directions, making a fixed angle \( \theta \) with the imposed field, along which the two \( \mu(\theta) \) surfaces touch.

18.7. The travel of wave-packets and the ray-direction

It is well known from the theory of physical optics that in a non-isotropic medium in which the phase velocity varies with direction a wave-packet will not, in general, travel along the wave-normal direction. Its direction of travel, known as the ray-direction, can be shown to be along the normal to the \( \mu(\theta) \) surface. In a magneto-ionic medium, also, the ray and wave-normal directions do not in general coincide, and it is possible to calculate the ray-direction from the \( \mu(\theta) \) curves discussed in §18.5. Some workers have, however, discussed the problem differently in terms of the Poynting vector, and energy flow, as follows.

The movement of a wave-packet corresponds to the flow of localised energy, so that the ray-direction is also the direction of the mean Poynting vector (see Scott (1950a)). Now in a characteristic wave the electric-field vector is such that, in general, its end will move round an ellipse in a plane inclined to the wave-front, whereas the magnetic-field vector lies always in the wave-front. The Poynting vector, which is, at each instant, perpendicular both to the electric and the magnetic fields, therefore moves round the surface of a cone with one edge perpendicular to the wave-front, and one perpendicular to the plane of the electric field, as indicated in fig. 18.5. The time-average direction of this vector gives the direction of mean energy flow, or the ray-direction. In the simpler problem of wave propagation in a crystal the Poynting vector is also inclined to the wave-front, but in a direction which remains fixed throughout the wave-cycle.

The time-average Poynting vector, for waves passing through a magneto-ionic medium, is found by taking half the real part of the complex Poynting vector,† as represented by \( \mathbf{E} \wedge \mathbf{H}^* \). Scott (1950a) has made detailed calculations of ray paths by this method.

† See Stratton (1941).
APPENDIX

To show that if \( \kappa = (v/2c)(1/\mu)\left\{X/(1 \pm Y_L)^3\right\} \) (A. 1)
and 
\[ \mu^2 = 1 - X/(1 + Y_L) \] (A. 2)
with 
\[ X = \omega^2 \nu / \omega^4, \quad (dX/d\omega) = -2X/\omega \]
and 
\[ Y_L = \omega_L / \omega, \quad (dY_L/d\omega) = -Y_L/\omega, \]
then 
\[ \kappa = (v/2c)(\mu' - \mu)(1 \pm \frac{1}{2} Y_L)^{-1}. \] (A. 3)

From (A. 2) 
\[ 2\mu(d\mu/d\omega) = \frac{2X}{\omega}(1 \pm \frac{1}{2} Y_L)^{2}, \]
hence 
\[ \omega(d\mu/d\omega) = (1/\mu)(1 \pm \frac{1}{2} Y_L)\left\{X/(1 \pm Y_L)^3\right\}. \] (A. 4)

But, from (11.1.2), 
\[ \omega(d\mu/d\omega) = \mu' - \mu, \] (A. 5)
and combination of (A. 4), (A. 5) and (A. 1) gives 
\[ \kappa = (v/2c)(\mu' - \mu)(1 - \frac{1}{2} Y_L)^{-1}. \] (A. 3)

Q.E.D.

BIBLIOGRAPHY

PART I

Annotated references arranged alphabetically under authors

APPLETON, E. V. (1927). U.R.S.I. Proc., Washington Assembly. [States the magneto-ionic equation in the symbols of Lorentz, but without proof.]


APPLETON, E. V. & BUILDER, G. (1933). The ionosphere as a doubly refracting medium. Proc. Phys. Soc. 45, 208. [Discusses magneto-ionic splitting and the importance of the quantity \( \omega_i \) of our §7.2.]

APPLETON, E. V. & CHAPMAN, F. W. (1932). The collisional friction experienced by vibrating electrons in ionised air. Proc. Phys. Soc. 44, 246. [The expression \( g = mv \) is derived for the friction coefficient in terms of collision frequency.]


Bibliography


Bailey, V. A. & Mattyn, D. F. (1934). The influence of electric waves on the ionosphere. *Phil. Mag.* 18, 369. [The original key paper on 'Ionospheric cross-modulation', 'Wave interaction', or 'The Luxembourg effect'.]


Banerji, R. B. (1952). On the origin of the third ionospheric echo. *Indian J. Phys.* 26, 28. [The third reflection point is said to occur when the group velocity is zero.]


Booher, H. G. (1932). The application of the magneto-ionic theory to the ionosphere. *Proc. Roy. Soc. A*, 150, 267. [Detailed consideration of the heights at which all frequencies would be reflected in a model ionosphere, both with and without inclusion of the Lorentz term. Also considers absorption. In this paper our $X$, $Y$, $Z$ are written $x$, $y$, $z$, and $X$, $Y$ have other meanings.]


Cowlings, T. G. (1943). The electrical conductivity of an ionized gas in a magnetic field, with applications to the solar atmosphere and the ionosphere. *Proc. Roy. Soc. A*, 183, 453. [Calculates the collision frequency of electrons with ions.]


BIBLIOGRAPHY


DAVIDS, N. (1953). Optic axes and critical coupling in the ionosphere. J. Geophys. Res. 58, 311. [The condition for critical coupling is that the direction of propagation should be along an optic axis.]


ECCLES, W. H. (1912). On the diurnal variation of the electric waves occurring in nature, and on the propagation of electric waves round the bend of the earth. Proc. Roy. Soc. A, 89, 79. [The first attempt to formulate mathematically the suggestions of Kennelly and Heaviside.]


FÖRSTERLING, K. (1942). Über die Ausbreitung elektromagnetischer Wellen in einem magnetisierten Medium bei senkrechter Inzidenz. Hochfrequenztech. u. Elektroakust. 59, 10. [Investigates the conditions under which one characteristic wave can give rise to the other.]


GHOSE, S. (1938). Dispersion, absorption, and polarisation curves for radio-wave propagation in the ionosphere. Indian J. Phys. 12, 341. [A few curves of p(\(X\)) and \(\chi(X)\) and polarisation for some special values of \(Y\) and \(\nu\)].

GINSBURG, V. (1944). The refractive index of an ionised gas. Wireless Engr. (Abstract), 22, 393. [Shows that Lorentz Term should not be included.]


GOUBAU, G. (1935a). Zur Dispersionstheorie der Ionosphäre. Hochfrequenz. u. Elektroakust. 45, 179. [Plots a series of curves to show \(p(X)\) and \(\chi(X)\) for different values of \(Y\) and \(\nu\). The Lorentz Term is omitted.]


HARTREE, D. R. (1931a). The propagation of electro-magnetic waves in a refracting medium in a magnetic field. Proc. Camb. Phil. Soc. 27, 143. [Develops the theory from a scattering point of view. Discusses the Lorentz Term and decides to include it.]


HEADING, J. & WHIPPLE, R. T. P. (1932). The oblique reflection of long wireless waves from the ionosphere at places where the earth's magnetic field is regarded as vertical. Phil. Trans. A, 244, 450.


HUXLEY, L. G. H. (1938). The propagation of electromagnetic waves in an ionised atmosphere. Phil. Mag. 25, 148. [Leads to conclusions which are stated to differ from those of Appleton in some respects.]


BIBLIOGRAPHY


Kelso, J. M. (1951b). Magneto-ionic calculations for 150 kc/s. Pen. State Coll. Tech. Rep. no. 24. [Tables and graphs of \( \mu(X), \chi(X) \) and \( R(X) \) for a series of different values of \( v \) with magnetic field appropriate to State College.]


Larmor, J. (1924a). Why wireless electric rays can bend round the earth. Nature, Lond., 114, 650. [Shows that if \( \nu < \omega \) refraction is possible but too much absorption.]

Larmor, J. (1924b). Why wireless electric rays can bend round the earth. Phil. Mag. 48, 1025. [Shows the importance of considering the dielectric rather than the conducting properties of the ionosphere.]


### Bibliography

**Ross, C. & Fluke, C. S.** (1953). Magneto-ionic calculations. Wright Air Dev. Centre Tech. Rep. pp. 53–96. [Extensive tables of \( \mu(X) \), \( \chi(X) \) and \( R(X) \), for a wide range of wave-frequency and collision-frequency, for vertical propagation at State College, Pennsylvania.]


**Shinn, D. H.** (1953). Tables of group refractive index for the Ordinary ray in the ionosphere. *Phys. Soc. Conf. Ionosph. p. 402. [Cover the range 0 to 1 in \( (1-X) \), \( \delta \) to 1 in \( Y, \theta \) to 80° in \( \theta \).]


PART II

Abbreviated references appropriate to sections of the book

CHAPTER 1

1-4 Appleton (1925), (1927), (1932), (1936), (1939)
Goldstein (1928)
Taylor & Hulbert (1926)
Larmor (1924a, b)
Eccles (1912)
Appleton & Barnett (1925)
Nichols & Schelleng (1925a, b)
Breit (1927)
Lassen (1927)
Hartree (1929), (1931a)
Lorentz (1900)
Darwin (1924)
Drude (1902)

Appleton & Chapman (1932)
Huxley (1938), (1940)
Jancel & Kahn (1955)
Pfister (1955)
Lorentz (1909)

CHAPTER 5

Ratcliffe (1956)

CHAPTER 6

Ratcliffe (1933)
Taylor (1933)

CHAPTER 7

Booker (1934)
Alden, de Bettencourt & Waterman (1950)
Ghose (1938)

2.2 Appleton & Builder (1933)
Booker (1934)

4.3

7-7

Booker (1934)

Baker & Green (1933)
Eckersley & Millington (1939)

CHAPTER 8

Booker (1934)
Whitehead (1952)

CHAPTER 9

9.1

Bailey & Somervile (1938)
Bailey (1934)

CHAPTER 10

Taylor (1934)
Goubau (1935a)
Westfold (1951)
Munck (1956)
Ghose (1938)
Rosa & Fluke (1953)
Kelso (1951b)
Martyn (1933)
Voogt (1948), (1949)

CHAPTER 11

Rai (1937)
Millington (1938)
Bajpai & Mathur (1937)
Goubau (1934a, b)
Kelso (1954)
Shin & Whale (1952)
Whale & Stanley (1950)
Shin (1955)
Mullaly (1956)

CHAPTER 12

Nicolet (1953)

12.2

12.4

Chapman (1931)

CHAPTER 13

Booker (1934), (1935)

13.2

13.4

Appleton & Builder (1933)
Appleton & Naismith (1932)
Appleton & Builder (1933)
Scott (1949)

13.5

Ellis (1953a, b), (1956)
Hogarth (1951)
Banerji (1952)
Dieminger & Möller (1949)
Landmark (1952)
Müller (1952)
Rydbeck (1950), (1951)
Scott (1948)
Eckersley (1950)
Becker (1951)

13.6

Appleton (1937)
Lepechinsky (1955)

13.8

Appleton, Farmer & Ratcliffe (1938)
Booker & Berkner (1938a, b)
Farmer & Ratcliffe (1935)
Martyn & Munro (1938a, b), (1939)
Goubau (1934a, b)

13.9

Bajpai & Mathur (1937)
Booker & Berkner (1938a, b)
Rai (1937)
Budden (1955b)

13.10

Lepechinsky (1956)
Landmark & Lied (1957)
Becker (1950)
Wilkes (1940)
Heading & Whipple (1952)

13.11

Becker (1950)

14.2

Bailey & Martyn (1934)
Huxley & Ratcliffe (1949)

Chapter 14

Bailey & Martyn (1934)
Huxley & Ratcliffe (1949)

Chapter 15

Goubau (1934a, b)
Kelso (1951a)
Lorentz (1900)
Martyn & Munro (1938a, b), (1939)
17.3
Booker (1936)
Försterling (1942)
Rydbeck (1951)
Eckersley (1950)
Davids (1953)

17.4
Lindquist (1953)
Kelso, Nearhoff, Nertney & Waynick (1951)
Nertney (1953)
Parkinson (1955)

17.5
Eckersley & Millington (1939)
Booker (1936)
Baker & Green (1933)
Budden (1952)

Chapter 18
18.2, 3 and 4
Saha, Banerjea & Guha (1947)
Saha & Banerjea (1945)
Arzehls (1942)
Banerjea (1947)
Westfold (1949)

18.5
Clemmow & Mullaly (1955)

18.6
Davids (1953)
Westfold (1949)
Turner (1954)

18.7
Scott (1950a)
Stratton (1941)

Chapter 19
17.2
Hartree (1931b)
Försterling & Wüster (1950)
Saha & Rai (1937)
Gans (1915)

INDEX
The references are to sections unless otherwise stated

Absorption coefficient, 4.4
deviative, 12.5(b), 13.7
energy transfer causing, 3.2, 3.3
in ionosphere, 13.5
macroscopic theory, ch. 4
microscopic theory, ch. 5
non-deviative, 12.5(a), 13.6
re-radiation causing, 5.4
Amplitude-phase diagram, 3.3
Anisotropic medium, ch. 18
Appleton’s equations
without collisions, 2.5
with collisions, 2.6
Graphs of: without collisions, ch. 6;
with collisions, ch. 10
Characteristic waves, 2.3, 2.5, 3.8
Collisions
average damping effect, 4.3
effect on equations, 2.6
effect on curves, ch. 9
Collision frequency
with neutral and charged particles, 4.2
in the ionosphere, 4.2, 14.2
of ions, 16.4
Complex refractive index, 4.4
Constitutive relations, 2.1
for magneto-ionic medium without collisions, 2.4
for magneto-ionic medium with collisions, 2.6, 18.2
Coupling
of wave, 13.5, 13.10, ch. 17, 17.2
echo, 17.3
critical, 18.6
between ions and electrons, 16.2
Critical coupling, 18.6
Critical frequency, 12.2
Crystal compared with magneto-ionic medium, ch. 18
Damping by collisions, 4.3
Deviative absorption, 12.5(b)
Dielectric constant, related to polarisability, 2.2
Dielectric tensor, 18.2, 18.3
Energy of electrons, 5.2, 14.4
Equivalent height, 12.4
Extraordinary wave, 1.6, 7.3
'Fourth reflection condition', 13.9
Full-wave theory, ch. 17
Graphs of Appleton’s equations
without collisions, ch. 6
with collisions, ch. 10
Group velocity, ch. 11, 13.1
Gyro-frequency splitting, 13.1, 13.8
Gyro-frequency of ions, 16.4
Height equivalent, 12.4
of reflection, 12.3
Ionospheric cross modulation, 14.2
Ions
effect on collisions, 4.2
effect on magneto-ionic splitting, 16.3
effect on propagation, ch. 16
Inhomogeneous media, ch. 17
Longitudinal case, no collisions, 6.2
Lorentz Term, 2.1, 2.3, ch. 15
Low frequencies, 13.10
Luxembourg effect, 14.2
Macroscopic theory, 2.1, 15.2
Magneto-ionic medium, 2.1
Magneto-ionic splitting, 13.1, 13.2
effect of collisions, 13.3
Maxwell’s equations, 2.2
Microscopic theory, 2.1, ch. 3, ch. 5
Mixture of ions and electrons, ch. 16
Model ionospheres, 8.4, 12.1, 12.4, ch. 13
Non-deviative absorption, 12.5(a)
Numerical computations, ch. 10
Numerical magnitudes of electrons’ motions, ch. 14
INDEX

Oblique reflection, 15.4(a)
Optic axes, 18.6
Orbits of electrons, 14.3
Ordinary wave, 6.6, 7.3
Penetration frequency, 12.2, 13.1
Phase, of scattered waves, 3.3
Polarisation
  of characteristic waves, 2.3, 2.5, 2.6, ch. 7
  emergent, 17.5
  of medium, 2.1, 2.3
  limiting, 7.3, 17.3, 17.5
Polarisability
  defined, 2.1
  related to refractive index, 2.2
tensor, 2.4, 18.2
Poynting vector, 5.2, 18.6
Principal directions, 18.4
Quasi-longitudinal (QL) approximation, ch. 8, 9.6, 13.6, 13.7
Quasi-transverse (QT) approximation, ch. 8, 9.7, 13.7
Rays, 18.7
Reflection
  conditions for, 9.4, 9.5, 10.2, 10.3, 17.2
  from effective conductor, 12.3
  height of, 12.3
  refractive index related to polarisability, 2.2
Refractive index surface, 18.5
Re-radiation
  causing absorption, 5.4
  causing change of velocity, ch. 3
Scattered waves
  as origin of altered velocity, 3.1, 3.2, 3.3, 3.4
  as origin of absorption, 5.4
Self-consistent wave in a dielectric, 3.5, 3.6, 3.7
Sellmeier theory, 15.2
Splitting
  effect of collisions on, 13.3
gyro-frequency, 13.1
  magneto-ionic, 13.1, 13.2
triple, 13.5
Susceptibility (see polarisability)

Tensor method of calculation, 18.2
Thin layer, reflection from, 17.1
Transient velocities and their part in absorption, 5.2, 5.3, 14.4
Transverse case, no collisions, 6.3
Triple splitting, 13.5
Tunnel effect, 17.1

Velocities of electrons, 14.2, 14.3
Very low frequencies, 13.10, 15.4

Wave
  in anisotropic medium, 2.3
  interaction, 14.2
  packets, 18.7
  in varying medium, 7.3, ch. 17
  velocity surface, 8.5
Whistlers, 15.4(c)

Zones, division of wave-front into, 3.3
Z-trace, 13.5
A COMPANION BOOK

RADIO WAVES IN THE IONOSPHERE

By K. G. BUDDEN

Fellow of St John's College and Lecturer in Physics in the University of Cambridge

This book gives the full mathematical theory of the propagation of radio waves in the ionosphere and their reflection from it. It is complementary to The Magneto-ionic Theory since Dr Budden gives the mathematical development of many topics mentioned by Mr Ratcliffe.

The first half is mainly concerned with the topics usually called ‘ray theory’, being applicable to homogeneous ionised media and to the reflection of high frequency radio waves at vertical and oblique incidence. The Appleton-Hartree formula and the Booker quartic equation are derived and fully discussed. The second half deals with ‘full wave theory’, which is used when the ray theory approximations fail. It is applied to the reflection of radio waves of very low frequency, and also includes detailed studies of the processes of reflection, coupling between characteristic waves, and partial penetration. There are chapters on the Airy integral function and the Stokes phenomenon, the phase integral method, numerical methods, and reciprocity.

The book will serve as a textbook for those comparatively new to the subject and as a reference book for practising engineers and research workers in the field of radio communication, for whom an understanding of the mathematical methods is important in solving practical problems.

CAMBRIDGE UNIVERSITY PRESS

Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York 22, N.Y.
West African Office: P.O. Box 33, Ibadan, Nigeria

PRINTED IN GREAT BRITAIN