4

Magnetoionic theory 1.
Polarisation and refractive index

4.1. Plane wave and homogeneous plasma

Magnetoionic theory is that branch of our subject that studies the properties of a plane electromagnetic wave in a homogeneous cold magnetoplasma, by using Maxwell's equations together with the electric permittivity tensor as given by the constitutive relation, ch.3. A plane wave was defined in § 2.8 and the z axis was there chosen to be parallel to the wave normal. This choice is used here. A progressive wave was defined as one in which the z dependence of all field components is only through the same factor $\exp(-\mathrm{i}knz)$ where $k=\omega/c$ and n is the refractive index. Thus for operations on any field component

$$\partial/\partial x \equiv 0, \quad \partial/\partial y \equiv 0, \quad \partial/\partial z \equiv -\mathrm{i}kn, \tag{4.1}$$

and hence Maxwell's equations in the form (2.45) give

$$nE_y = -\mathscr{H}_x \tag{4.2}$$

$$n\mathscr{H}_y = \varepsilon_0^{-1} D_x \tag{4.3}$$

$$nE_x = \mathscr{H}_y \tag{4.4}$$

$$n\mathscr{H}_x = -\varepsilon_0^{-1} D_y \tag{4.5}$$

$$\mathscr{H}_z = 0 \tag{4.6}$$

$$D_z = 0 \tag{4.7}$$

In later chapters a different choice of coordinates x, y, z is used so that the field components in (4.2)–(4.7) there have different meanings.

Equations (4.6)–(4.7) show that $\mathscr{H}$ and $\boldsymbol{D}$ are perpendicular to the wave normal. This must always apply to any plane wave. The constitutive relation gives $\boldsymbol{D}$ in terms of $\boldsymbol{E}$ thus

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} = \varepsilon_0 \boldsymbol{\varepsilon} \boldsymbol{E}. \tag{4.8}$$

This and equations (4.2)–(4.7) are the starting-point for the study of magnetoionic theory in this and the following chapter.

4.2. Isotropic plasma

For an isotropic plasma $\boldsymbol{\varepsilon}$ is a scalar ε and $\boldsymbol{D}$ is parallel to $\boldsymbol{E}$. Then the ratio $\mathscr{H}_x/E_y$ may be eliminated from (4.2), (4.5) to give

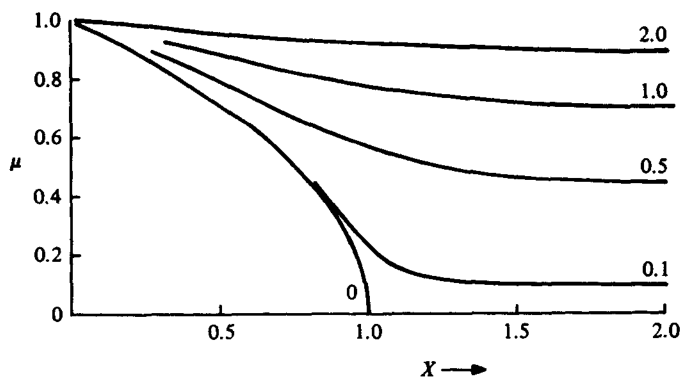
$$n^2 = \varepsilon = 1 - X/U \tag{4.9}$$

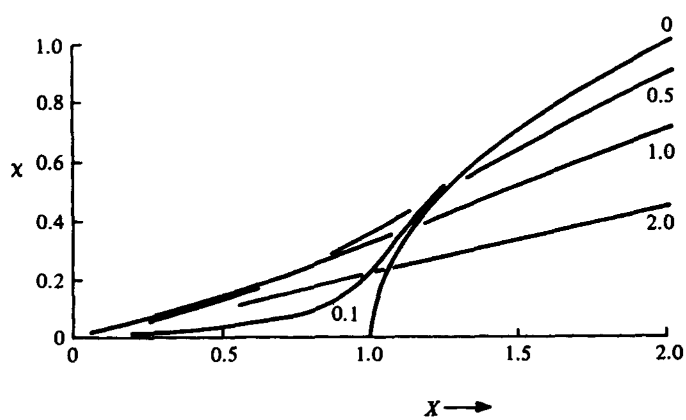
where (3.17) has been used. The same result is obtained by eliminating $\mathscr{H}_y/E_x$ from (4.3), (4.4). These two pairs of equations are independent of each other, which shows that the two linearly polarised plane waves, with their electric fields parallel to the y and x axes respectively, are propagated independently and (4.9) applies to both. They therefore travel with the same wave velocity $v = c/n$. Two such linearly polarised waves with any amplitudes and relative phase could be combined to give a wave which would in general be elliptically polarised.

For a collisionless plasma (4.9) gives

$$n^2 = 1 - X = 1 - f_N^2/f^2. \tag{4.10}$$

Fig. 4.1. Shows how the real part μ and the imaginary part $-\chi$ of the refractive index $n = \mu - i\chi$ depend on X for an isotropic plasma. The numbers by the curves are the values of $Z = \nu/\omega$.





This is the same as ε_3, (3.58), for a collisionless electron magnetoplasma. Its frequency dependence is shown in fig. 3.1. Thus if $f > f_N$, n is real and the wave is propagated without attenuation. If $f < f_N$, n is purely imaginary and the wave is evanescent; see § 2.14. The plasma frequency f_N is the cut-off frequency. The dispersion law (4.10) is the same as for one mode in a loss-free wave guide.

If collisions are included, (4.9) shows that n is complex. We choose the value with positive real part, and the imaginary part must then be negative. The wave travels in the direction of z increasing and is attenuated as it travels; see § 2.13. Fig. 4.1 shows how the real and imaginary parts of n depend on X for several different values of Z. More extensive curves of this kind and a full discussion are given by Ratcliffe (1959). It is clear that if Z is non-zero and real, there is no real value of X that makes $n^2 = 0$. In later chapters, however, it will prove useful to assume that X and Z are analytic functions of the height h (in later chapters denoted by z.) They must be real and positive when h is real, but when h is complex, X and Z can take complex values. Then n^2 can be zero for one or more complex values of h. These zeros of n^2 in the complex h plane play a very important part in the theory of the reflection of radio waves.

4.3. Anisotropic plasma. The wave polarisation

In an anisotropic plasma D_y cannot be expressed in terms of E_y alone, and D_x cannot be expressed in terms of E_x alone. Hence the pairs of equations (4.2), (4.5) and (4.3), (4.4) are no longer independent. This leads to a restriction on the possible states of polarisation of the wave, as will now be shown.

The z axis is chosen parallel to the wave normal but the x and y axes may be rotated about the z axis, so that we are free to impose a further condition. The axes are therefore chosen so that the vector $\boldsymbol{Y}$ is in the x-z plane at an angle Θ to the z axis, and

$$l_x = \sin\Theta, \quad l_z = \cos\Theta \tag{4.11}$$

have the same sign, so that $\tan\Theta$ is positive. Then $l_y = 0$ and the constitutive relation (3.25) gives

$$-\varepsilon_0 X\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} = \begin{pmatrix} U & \mathrm{i}Y\cos\Theta & 0 \\ -\mathrm{i}Y\cos\Theta & U & \mathrm{i}Y\sin\Theta \\ 0 & -\mathrm{i}Y\sin\Theta & U \end{pmatrix}\begin{pmatrix} P_x \\ P_y \\ P_z \end{pmatrix}. \tag{4.12}$$

For the more general plasma, these same axes were used to give the form (3.54) of the permittivity tensor.

With this choice of axes, the wave polarisation ρ of a progressive wave is defined thus

$$\rho = E_y/E_x. \tag{4.13}$$

Equations (4.2), (4.4) then show that

$$\rho = -\mathscr{H}_x/\mathscr{H}_y. \tag{4.14}$$

The component E_z may not be zero, but it does not enter into the definition of ρ. The component $\mathscr{H}_z$ is always zero.

The wave polarisation ρ is a complex number which shows how the transverse component of $\boldsymbol{E}$ varies with time. For example if ρ is real, E_x and E_y have the same phase and the transverse component of $\boldsymbol{E}$ is always parallel to a fixed line; this is called linear polarisation. If $\rho = \mathrm{i}$, E_x and E_y are in quadrature and have equal amplitudes, and the polarisation is circular. To an observer looking in the direction of the wave normal, that is in the direction of positive z, the electric vector observed at a fixed point would appear to rotate anticlockwise. This is called left-handed circular polarisation. If $\rho = -\mathrm{i}$ the rotation is clockwise, and the polarisation is called right-handed circular polarisation. These conventions are used in some recent physics text books, and are followed in this book. They accord with the recommendations of the Institute of Electrical and Electronics Engineers (1969), and of the International Astronomical Union (1973). Some authors use different conventions, so it is important to state which system is being used. If ρ is complex, the polarisation is elliptical, and (4.13), (4.14) show that the magnetic vector $\mathscr{H}$ and the transverse component of $\boldsymbol{E}$ traverse similar ellipses in the same sense. The ellipses for $\boldsymbol{E}$ and for $\mathscr{H}$ have their major axes at right angles. There are alternative ways of describing the wave polarisation; see § 4.5. For a full discussion see Rawer and Suchy (1967, § 7).

We shall now apply (4.13), (4.14) to an electron plasma and use the form (4.12) of the constitutive relation. Equations (4.7), (4.8) show that $\varepsilon_0 E_z + P_z = 0$. If this is used in the z component of (4.12) it gives

$$(U - X)P_z = \mathrm{i}\,Y \sin\Theta\, P_y. \tag{4.15}$$

Now divide (4.5) by (4.3) and use (4.13), (4.14) and (4.8). This gives

$$\frac{D_y}{D_x} = -\frac{\mathscr{H}_x}{\mathscr{H}_y} = \frac{E_y}{E_x} = \frac{P_y}{P_x} = \rho. \tag{4.16}$$

The x and y components of (4.12) give

$$-\varepsilon_0 X E_x = U P_x + \mathrm{i}\,Y \cos\Theta\, P_y \tag{4.17}$$

$$-\varepsilon_0 X E_y = -\mathrm{i}\,Y\cos\Theta\, P_x + \left(U - \frac{Y^2 \sin^2\Theta}{U - X}\right) P_y \tag{4.18}$$

where (4.15) has been used in (4.18). Now divide (4.18) by (4.17) and use (4.16). Then

$$\rho = \frac{-\mathrm{i}\,Y\cos\Theta + \rho\{U - Y^2\sin^2\Theta/(U - X)\}}{U + \mathrm{i}\,Y\rho\cos\Theta} \tag{4.19}$$

whence

$$\rho^2 - i\rho\frac{Y\sin^2\Theta}{(U-X)\cos\Theta} + 1 = 0. \tag{4.20}$$

This is a quadratic equation for ρ and shows that, in a homogeneous electron magnetoplasma, with ν independent of electron velocity, a progressive wave must have a polarisation given by one of the two solutions

$$\rho = \frac{\frac{1}{2}iY\sin^2\Theta \pm i\{\frac{1}{4}Y^2\sin^4\Theta + \cos^2\Theta(U-X)^2\}^{\frac{1}{2}}}{(U-X)\cos\Theta}. \tag{4.21}$$

Equation (4.20) is the polarisation equation of magnetoionic theory and plays a most important part in the theory of later chapters.

If a plane wave is present with its wave normal parallel to the z axis and with any given polarisation, it must be resolved into two component waves with the two polarisations (4.21). These waves in general have different refractive indices (see (4.46) below) so that the composite wave is not a progressive wave because its fields do not have a unique z dependence satisfying (4.1).

If $X = 0$, the medium is free space. Yet (4.21) shows that apparently ρ can have one of two values only. But now (4.46) shows that both refractive indices are unity. A wave with any desired polarisation can be resolved into components with the polarisations (4.21), but these both travel with wave velocity c and the resultant polarisation does not change. The composite wave remains a progressive wave. This property is important in the theory of the limiting polarisation for a wave emerging from the ionosphere; see §§ 17.10, 17.11.

For a more general type of plasma, with ions allowed for or with a velocity dependent collision frequency, or both, we use the form (3.54) for the permittivity tensor with double subscripts x, y, z to indicate its components. Then (4.7) gives

$$E_z = -(\varepsilon_{zx}E_x + \varepsilon_{zy}E_y)/\varepsilon_{zz}. \tag{4.22}$$

This replaces (4.15). Division of D_y by D_x gives, from (4.16) with (4.22) after some reduction

$$\rho = \frac{\varepsilon_{zz}\varepsilon_{yx} - \varepsilon_{yz}\varepsilon_{zx} + (\varepsilon_{yy}\varepsilon_{zz} - \varepsilon_{yz}\varepsilon_{zy})\rho}{\varepsilon_{xx}\varepsilon_{zz} - \varepsilon_{xz}\varepsilon_{zx} + (\varepsilon_{xy}\varepsilon_{zz} - \varepsilon_{xz}\varepsilon_{zy})\rho}. \tag{4.23}$$

The elements of $\boldsymbol{\varepsilon}$ from (3.54) are now used, and J from (3.51). Then (4.23) is the quadratic equation

$$\rho^2 - i\rho\frac{2J\sin^2\Theta}{\varepsilon_3(\varepsilon_1 - \varepsilon_2)\cos\Theta} + 1 = 0. \tag{4.24}$$

This applies for the general cold magnetoplasma, but if the special values (3.50) are used, it is the same as (4.20).

4.4. Properties of the polarisation equation

From (4.21) it is clear that when collisions are neglected, so that $U = 1$, both values of ρ are purely imaginary. The two polarisation ellipses then have their major axes parallel to either the x or the y axis.

The condition that two solutions of the quadratic (4.20) shall be equal is

$$\tfrac{1}{4}Y^2\sin^4\Theta + \cos^2\Theta(U - X)^2 = 0. \tag{4.25}$$

When X and Z are real, this is only possible if

$$X = 1 \text{ and } Z = Z_t \tag{4.26}$$

where

$$Z_t = \tfrac{1}{2}Y\sin^2\Theta/|\cos\Theta| \tag{4.27}$$

or equivalently

$$\nu_t = \omega_t \quad \text{where} \quad \omega_t = \tfrac{1}{2}\omega_H\sin^2\Theta/|\cos\Theta|. \tag{4.28}$$

Thus ω_t is a 'transition' value of the collision frequency, and is independent of the wave frequency. Many authors, including Budden (1961a), have called it the critical frequency and used for it the symbol ω_c, but the word 'critical' is now used for too many different things in physics. The transition frequency ω_t depends only on the strength of the earth's magnetic field and its inclination $\pi - \Theta$ to the wave normal. Suppose that $\cos\Theta$ is positive, as it is in the northern hemisphere for a radio wave with its wave normal vertically upwards. Then, if the condition (4.26) holds, both values of ρ are -1, so that, for both polarisations, the electric field is in the same plane at $-45°$ to the x axis. In these conditions it is possible, however, for another wave to exist whose electric field has a component at right angles to this plane; see §4.14. In the southern hemisphere, when (4.26) is true, both values of ρ are $+1$; see fig. 4.2.

Let the two values (4.21) of ρ be ρ_O and ρ_E. The subscripts O and E mean 'ordinary' and 'extraordinary', respectively. The meanings of these terms are discussed in the following section and in §§4.11, 4.15, 5.6. Now (4.20) shows that

$$\rho_O\rho_E = 1. \tag{4.29}$$

Choose new axes x', y', z', formed from x, y, z by rotation through an angle ψ about the z axis and let E'_x, E'_y, ρ'_O, ρ'_E be the new values of E_x, E_y, ρ_O, ρ_E. Then for a wave with $E_y = \rho_O E_x$

$$E'_x = E_x(\cos\psi + \rho_O\sin\psi), \quad E'_y = E_x(-\sin\psi + \rho_O\cos\psi) \tag{4.30}$$

whence

$$\rho'_O = (\rho_O - \tan\psi)/(1 + \rho_O\tan\psi). \tag{4.31}$$

Similarly it can be shown that

$$\rho'_E = (\rho_E - \tan\psi)/(1 + \rho_E\tan\psi). \tag{4.32}$$

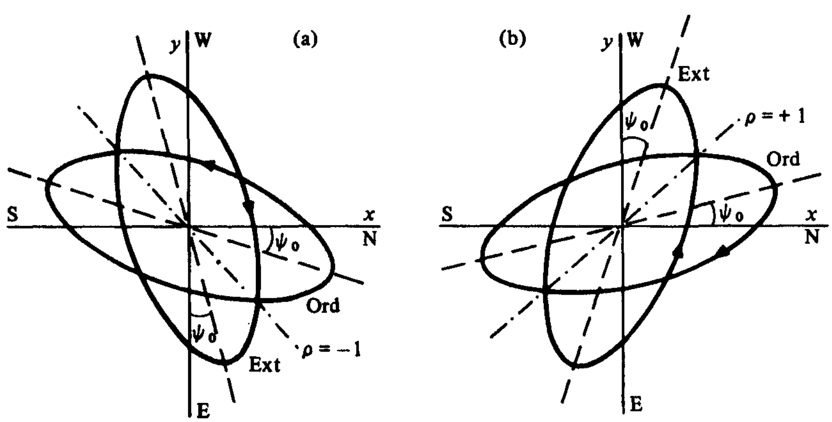
Now (4.29), (4.20) show that if $\psi = \frac{1}{4}\pi$

$$-\rho'_{\mathrm{O}} = \rho'_{\mathrm{E}} = \frac{\mathrm{i}\{\frac{1}{4}Y^2 \sin^4\Theta + (U-X)^2\cos^2\Theta\}^{\frac{1}{2}}}{\frac{1}{2}\mathrm{i}Y\sin^2\Theta + (U-X)\cos\Theta} \tag{4.33}$$

where the square root has a positive real part when $Z < Z_t$, (4.27). Hence the two polarisation ellipses are mirror images of each other in a plane at 45° to the original axes, that is to the magnetic meridian plane. The ellipses are identical in shape, and their field vectors have opposite directions of rotation. The two major axes are in the same quadrant; see fig. 4.2.

The polarisation ρ is defined by (4.13) in terms of a right-handed system of axes in which the z axis is the wave normal. For a wave travelling in the opposite direction, the z axis and one of the other axes must be reversed. To ensure that $\tan\Theta$ remains positive (§ 4.3 and (4.11)), this other axis must be the x axis. The middle term of (4.20) then changes sign so that both values of ρ change sign. There can be two waves with anti-parallel wave normals and ρ values differing only in sign. The polarisation ellipses for these waves are exactly the same. The ρs have opposite signs only because of the change of axes. This convention is used in this and the following chapter. In some problems discussed in later chapters, waves in both directions are present, but it is convenient to define ρ in terms of the *same* set of axes for both waves. Then, for ordinary waves travelling in either direction, ρ has the same value ρ_{O}, and similarly for extraordinary waves.

Fig. 4.2. The two polarisation ellipses of the horizontal component of the electric field for waves with vertical wave normals, as seen looking downwards in the ionosphere. (a) is for the northern hemisphere and (b) for the southern, and N E S W denote the magnetic cardinal points. The arrows show the sense of rotation when $X < 1$, and should be reversed for $X > 1$. The ellipses for the magnetic field are obtained by rotating these figures through 90°.



4.5. Alternative measure of the polarisation. Axis ratio and tilt angle

Let
$$\rho_O = \tan\gamma = u + \mathrm{i}v \tag{4.34}$$
and
$$\gamma = \psi_O + \mathrm{i}\phi_O \tag{4.35}$$
where u, v, ψ_O, ϕ_O are real. The value ρ'_O (4.31), referred to new axes as in the preceding section is then
$$\rho'_O = \tan(\gamma - \psi) \tag{4.36}$$
which shows that ρ'_O is purely imaginary when $\psi = \psi_O$ or $\psi = \psi_O + \frac{1}{2}\pi$. These are the inclinations to the x axis of the major and minor axes respectively of the polarisation ellipse for the ordinary wave. Similarly the polarisation ellipse of the extraordinary wave has its major and minor axes at angles $\pm\frac{1}{2}\pi - \psi_O$ and $-\psi_O$ to the x axis. Substitution of (4.34) in the quadratic (4.20) gives
$$\sin 2\gamma = -2\mathrm{i}\,Y^{-1}(1 - X - \mathrm{i}Z)\cos\Theta/\sin^2\Theta \tag{4.37}$$
whence
$$\sin 2\psi_O \cosh 2\phi_O = -2Y^{-1}Z\cos\Theta/\sin^2\Theta, \tag{4.38}$$
$$\cos 2\psi_O \sinh 2\phi_O = -2Y^{-1}(1-X)\cos\Theta/\sin^2\Theta. \tag{4.39}$$
This gives two values of ψ_O and we may choose the value for which $|\psi_O| \leqslant \frac{1}{2}\pi$. If $\cos\Theta$ is changed but does not become zero or change sign, (4.38) shows that ψ_O does not change sign, and (4.39) shows that, if X is kept constant, $|\psi_O|$ never attains the value $\frac{1}{4}\pi$. It is shown in §4.11 that for transverse propagation, $\cos\Theta = 0$, the ordinary wave is linearly polarised with $\boldsymbol{E}$ parallel to the x axis. The polarisation ellipse for $\boldsymbol{E}$ has degenerated to a line, and its major axis is parallel to the x axis. But (4.38) shows that now $\psi_O = 0$. Thus ψ_O is the direction of the major axis of the $\boldsymbol{E}$ ellipse for the ordinary wave. When $\cos\Theta$ is small, $|\psi_O| < \frac{1}{4}\pi$ and this remains true for all Θ. The angle ψ_O is called the 'tilt angle' for the ordinary wave. It is the angle between the magnetic meridian and the major axis of the polarisation ellipse for the component of $\boldsymbol{E}$ perpendicular to the wave normal.

Let ψ in (4.36) take the value ψ_O. Then ρ'_O is purely imaginary and its modulus is the ratio: minor axis/major axis. This is called the 'axis ratio'. Its value, from (4.34), (4.35) is $|\tan\mathrm{i}\phi_O| = |\tanh\phi_O|$, and it can be shown from (4.34) that
$$\tanh 2\phi_O = 2v/(1 + u^2 + v^2). \tag{4.40}$$
The concepts of 'axis ratio' and 'tilt angle' were used for radio waves by Kelso *et al.* (1951).

Consider a radio wave travelling vertically upwards in the ionosphere in the northern hemisphere. Here the vertical component of the earth's magnetic field is downwards, so that the vertical component of $\boldsymbol{Y}$ is upwards, and $\cos\Theta$ is positive. Then (4.38) shows that ψ_O is negative. Thus, where it is permissible to treat the ionosphere as a homogeneous electron plasma, the major axes of the polarisation ellipses for electric field of both waves must be in the quadrant between magnetic north and magnetic east; see fig. 4.2(a). No progressive wave is possible with these major axes in the north-west and south-east quadrants. This applies also for vertically downgoing waves because the polarisation ellipses are the same as for upgoing waves; see end of §4.4. In the southern hemisphere $\cos\Theta$ is negative for upgoing waves. The major axes for the $\boldsymbol{E}$ ellipses are here always in the north-west and south-east quadrants. This is illustrated in fig. 4.2(b).

A more detailed discussion of the wave polarisation of a downcoming radio wave is given in §§17.10, 17.11.

For the more general type of plasma, when the quadratic (4.24) is used, the result (4.37) becomes

$$\sin 2\gamma = -2\mathrm{i}\frac{\varepsilon_3(\varepsilon_1-\varepsilon_2)\cos\Theta}{J\sin^2\Theta} \tag{4.41}$$

whence the axis ratio and tilt angle can be found exactly as described above.

4.6. Refractive index 1. The dispersion relation

If $\mathscr{H}_y$ is eliminated from (4.3), (4.4), they give

$$D_x = \varepsilon_0 n^2 E_x \tag{4.42}$$

and similarly (4.2), (4.5) give

$$D_y = \varepsilon_0 n^2 E_y. \tag{4.43}$$

These equations are the same as would be obtained for an isotropic dielectric, in which case n^2, the square of the refractive index, is equal to the scalar electric permittivity. This result may not, however, have any particular physical significance. It would not be true if the magnetic permeability differed from unity, and it is not true for the z components because $D_z = 0$ but in general $E_z \neq 0$.

From (4.8) with (4.42), (4.43)

$$P_x = \varepsilon_0(n^2-1)E_x, \quad P_y = \varepsilon_0(n^2-1)E_y. \tag{4.44}$$

From this substitute E_x in (4.17) for the electron plasma, and divide by P_x. Since $P_y/P_x = \rho$ from (4.16), the result is

$$X/(n^2-1) = -U - \mathrm{i}\rho Y\cos\Theta \tag{4.45}$$

whence

$$n^2 = 1 - \frac{X}{U + \mathrm{i}\rho Y\cos\Theta}. \tag{4.46}$$

Now substitute for ρ from (4.21). This gives

$$n^2 = 1 - \frac{X(U-X)}{U(U-X) - \frac{1}{2}Y^2\sin^2\Theta \pm \{\frac{1}{4}Y^4\sin^4\Theta + Y^2\cos^2\Theta(U-X)^2\}^{\frac{1}{2}}}. \tag{4.47}$$

This formula, for the square of the two refractive indices of a cold electron magnetoplasma, with the electron collision frequency independent of velocity, has for many years been known as the Appleton–Hartree formula. This form, apart from the notation, is essentially the same as Appleton's (1932) form and the method of deriving it given above is almost the same as Appleton's method. Both Appleton (1932) and Hartree (1931) included the Lorentz polarisation term (§ 3.3), but Appleton also gave a version without it. It is now known that the Lorentz term should be omitted; see § 3.3. A version of (4.47) without it was given by Goldstein (1928). An equivalent formula to (4.47), also without the Lorentz term, was derived by Lassen (1927) who used a somewhat different method. Some authors now call (4.47), or its equivalent, the Appleton–Lassen formula, and this name seems to be coming into increasing use; see Rawer and Suchy (1976). The formula (4.47) is often needed for a collisionless plasma so that $U = 1$. The result is given here for future reference:

$$n^2 = 1 - \frac{X(1-X)}{1 - X - \frac{1}{2}Y^2\sin^2\Theta \pm \{\frac{1}{4}Y^4\sin^4\Theta + Y^2\cos^2\Theta(1-X)^2\}^{\frac{1}{2}}}. \tag{4.48}$$

For the more general plasma, (4.17) cannot be used, and instead we substitute E_z, from (4.22), into the D_x element of (4.8). With (4.42) and $E_y/E_x = \rho$, this gives

$$n^2 = [(\varepsilon_{zz}\varepsilon_{xx} - \varepsilon_{xz}\varepsilon_{zx}) + \rho(\varepsilon_{zz}\varepsilon_{xy} - \varepsilon_{xz}\varepsilon_{zy})]/\varepsilon_{zz} \tag{4.49}$$

and on using the form (3.54) for the elements of $\boldsymbol{\varepsilon}$, it becomes

$$n^2 = \frac{\varepsilon_3(\varepsilon_1 + \varepsilon_2) + \mathrm{i}\rho\cos\Theta\,\varepsilon_3(\varepsilon_1 - \varepsilon_2)}{(\varepsilon_1 + \varepsilon_2)\sin^2\Theta + 2\varepsilon_3\cos^2\Theta}. \tag{4.50}$$

Then on substituting the solutions of (4.24) for ρ we obtain finally

$$n^2 = \frac{\varepsilon_1\varepsilon_2\sin^2\Theta + \frac{1}{2}\varepsilon_3(\varepsilon_1 + \varepsilon_2)(1 + \cos^2\Theta) \pm \{J^2\sin^4\Theta + \varepsilon_3^2(\varepsilon_1 - \varepsilon_2)^2\cos^2\Theta\}^{\frac{1}{2}}}{(\varepsilon_1 + \varepsilon_2)\sin^2\Theta + 2\varepsilon_3\cos^2\Theta}. \tag{4.51}$$

4.7. Longitudinal component of electric polarisation and electric field

Equation (4.7) shows that $D_z = 0$, but E_z and P_z in general are not zero. The components E_z, P_z are parallel to the wave normal and are called longitudinal components. For the electron plasma, from (4.15), (4.16) and (4.20)

$$P_z/P_x = \rho P_z/P_y = \rho\mathrm{i}y\sin\Theta/(U - X) = (\rho^2 + 1)\cot\Theta. \tag{4.52}$$

Thus the components of the complex vector $\boldsymbol{P}$ are in the ratio

$$P_x{:}P_y{:}P_z \propto U - X{:}\rho(U - X){:}\mathrm{i}Y\rho\sin\Theta \propto 1{:}\rho{:}(\rho^2 + 1)\cot\Theta. \tag{4.53}$$

If collisions are neglected, ρ is purely imaginary. Then P_x and P_z are in phase or antiphase and P_y is in quadrature with them. The normal to the polarisation ellipse for $\boldsymbol{P}$ is coplanar with $\boldsymbol{Y}$ and the wave normal. The vector $\boldsymbol{P}$ is of interest because it is proportional to the electron displacements caused by the wave.

From (4.52), with (4.7) and (4.44)

$$E_z/E_x = \rho E_z/E_y = -\rho \mathrm{i} Y \sin\Theta\,(n^2-1)/(U-X) = -\cot\Theta(\rho^2+1)(n^2-1). \tag{4.54}$$

If collisions are neglected so that n^2 is real, E_z and E_x are in phase or antiphase and E_y is in quadrature with them. The normal to the polarisation ellipse for $\boldsymbol{E}$ is coplanar with $\boldsymbol{Y}$ and the wave normal.

Equation (4.54) shows that the longitudinal component E_z of the electric field is zero only if either $\sin\Theta = 0$ (equivalently $\rho = \pm\mathrm{i}$), or $\rho = 0$, or $n^2 = 1$. The first of these conditions holds when the wave normal is parallel or antiparallel to the earth's magnetic field, and this is called longitudinal propagation. The second, $\rho = 0$, requires that $\cot\Theta = 0$ from (4.24), so that the earth's magnetic field is perpendicular to the wave normal, and this is called transverse propagation. The earth's field is also parallel to the electric field $\boldsymbol{E}$ of the wave. Then the only motion imparted to the electrons is parallel to the earth's magnetic field, and the electron motion is unaffected by the field. The third condition $n^2 = 1$ applies only when $X = 0$, that is for free space. It can be shown that one value of n^2 is unity when $X = 1$, $Z = 0$, but in this case ρ is infinite and E_z in (4.54) is not zero.

4.8. The flow of energy for a progressive wave in a magnetoplasma

It is interesting to evaluate the time average value $\boldsymbol{\Pi}_{\mathrm{av}}$ of the Poynting vector, given by (2.63), for a progressive wave in an electron magnetoplasma, so as to find the flux of energy. Since $\mathscr{H}_z = 0$, the coponents of $\boldsymbol{\Pi}_{\mathrm{av}}$ are

$$\Pi_x = \tfrac{1}{4}Z_0^{-1}(E_z\mathscr{H}_y^* + E_z^*\mathscr{H}_y), \quad \Pi_y = \tfrac{1}{4}Z_0^{-1}(E_z\mathscr{H}_x^* + E_z^*\mathscr{H}_x),$$
$$\Pi_z = \tfrac{1}{4}Z_0^{-1}(E_x\mathscr{H}_y^* + E_x^*\mathscr{H}_y - E_y\mathscr{H}_x^* - E_y^*\mathscr{H}_x). \tag{4.55}$$

In an isotropic medium E_z also is zero so that Π_x, Π_y are zero and the energy flow is in the direction of the wave normal, but this is not in general true for an anisotropic medium. It is convenient to express the components of $\boldsymbol{E}$ and $\mathscr{H}$ in terms of E_x by using (4.2), (4.4), (4.13), (4.14) and (4.54). Hence

$$\left.\begin{aligned}
\Pi_x &= -\frac{\mathrm{i}Y\sin\Theta}{4Z_0}\left\{\frac{n\rho^*(n^{*2}-1)}{U^*-X} - \frac{n^*\rho(n^2-1)}{U-X}\right\}|E_x|^2,\\
\Pi_y &= \frac{\mathrm{i}Y\sin\Theta\,\rho\rho^*}{4Z_0}\left\{\frac{(n^2-1)n^*}{U-X} - \frac{(n^{*2}-1)n}{U^*-X}\right\}|E_x|^2,\\
\Pi_z &= \frac{1}{4Z_0}(n+n^*)(1+\rho\rho^*)|E_x|^2.
\end{aligned}\right\} \tag{4.56}$$

These give the magnitude and direction of $\boldsymbol{\Pi}_{\mathrm{av}}$ when collisions are allowed for, and the values of n and ρ for either of the two characteristic waves may be inserted. The expressions are of interest for the special case when collisions are neglected so that n^2 is real, ρ is purely imaginary, and $U = U^* = 1$. Then $\rho^* = -\rho$. There are two important cases. The first is when n is real so that $n^* = n$. Then

$$\Pi_x = \frac{\mathrm{i}\rho n(n^2-1)Y\sin\Theta}{2Z_0(1-X)}|E_x|^2, \quad \Pi_y = 0, \quad \Pi_z = \frac{n}{2Z_0}(1-\rho^2)|E_x|^2, \tag{4.57}$$

which show that $\mathbf{\Pi}_{\mathrm{av}}$ is in the plane containing the earth's magnetic field and the wave normal, and makes an angle

$$\arctan\left\{\frac{\mathrm{i}Y\rho\sin\Theta(n^2-1)}{(1-X)(1-\rho^2)}\right\} = \arctan\left[\pm\frac{\tfrac{1}{2}Y\sin\Theta\cos\Theta\,(n^2-1)}{\{\tfrac{1}{4}Y^2\sin^4\Theta + (1-X)^2\cos^2\Theta\}^{\frac{1}{2}}}\right] \tag{4.58}$$

with the wave normal. Here ρ and $1/\rho$ from (4.21) have been used. The + sign applies for the ordinary wave and the − sign for the extraordinary wave. This result is derived by a different method in § 5.3 and its significance is discussed there.

The second important case is when n is purely imaginary, so that the wave is evanescent, and $n^* = -n$. Then (4.56) gives

$$\Pi_x = \Pi_z = 0, \quad \Pi_y = \frac{\mathrm{i}n\rho^2 Y\sin\Theta\,(n^2-1)}{2Z_0(1-X)}|E_x|^2 \tag{4.59}$$

which shows that $\mathbf{\Pi}_{\mathrm{av}}$ is perpendicular both to the wave normal and to the earth's magnetic field.

4.9. Refractive index 2. Alternative derivations and formulae

The refractive index n can be derived directly without going through the intermediate step of finding the polarisation ρ. We use (4.42), (4.43), (4.7) which come from Maxwell's equations. They are substituted in the constitutive relation (4.8) to give

$$\left.\begin{aligned}(\varepsilon_{xx}-n^2)E_x + \varepsilon_{xy}E_y + \varepsilon_{xz}E_z &= 0,\\ \varepsilon_{yx}E_x + (\varepsilon_{yy}-n^2)E_y + \varepsilon_{yz}E_z &= 0,\\ \varepsilon_{zx}E_x + \varepsilon_{zy}E_y + \varepsilon_{zz}E_z &= 0.\end{aligned}\right\} \tag{4.60}$$

For these three equations to have a solution, it is necessary that the determinant of the coefficients of E_x, E_y, E_z shall be zero. The form (3.54) for the elements of $\boldsymbol{\varepsilon}$ is now used. This gives

$$\begin{vmatrix} \tfrac{1}{2}(\varepsilon_1+\varepsilon_2)\cos^2\Theta + \varepsilon_3\sin^2\Theta - n^2 & \tfrac{1}{2}\mathrm{i}(\varepsilon_1-\varepsilon_2)\cos\Theta & G\sin\Theta\cos\Theta \\ -\tfrac{1}{2}\mathrm{i}(\varepsilon_1-\varepsilon_2)\cos\Theta & \tfrac{1}{2}(\varepsilon_1+\varepsilon_2) - n^2 & -\tfrac{1}{2}\mathrm{i}(\varepsilon_1-\varepsilon_2)\sin\Theta \\ G\sin\Theta\cos\Theta & \tfrac{1}{2}\mathrm{i}(\varepsilon_1-\varepsilon_2)\sin\Theta & \tfrac{1}{2}(\varepsilon_1+\varepsilon_2)\sin^2\Theta + \varepsilon_3\cos^2\Theta \end{vmatrix} = 0 \tag{4.61}$$

and on multiplying out, it gives a quadratic equation for n^2 which can be written in the various forms:

$$n^4\{\tfrac{1}{2}(\varepsilon_1+\varepsilon_2)\sin^2\Theta + \varepsilon_3\cos^2\Theta\} - n^2\{\varepsilon_1\varepsilon_2\sin^2\Theta + \tfrac{1}{2}\varepsilon_3(\varepsilon_1+\varepsilon_2)(1+\cos^2\Theta)\} + \varepsilon_1\varepsilon_2\varepsilon_3 = 0, \tag{4.62}$$

$$\varepsilon_3(n^2-\varepsilon_1)(n^2-\varepsilon_2)+n^2\sin^2\Theta(Gn^2+J)=0, \tag{4.63}$$

$$\tfrac{1}{2}(\varepsilon_1+\varepsilon_2)(n^2-\varepsilon_3)\left(n^2-\frac{2\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}\right)-n^2\cos^2\Theta(Gn^2+J)=0. \tag{4.64}$$

If the values (3.50) for the electron plasma are used this becomes

$$An^4-2Bn^2+C=0 \tag{4.65}$$

where

$$\left.\begin{aligned} A&=U^2(U-X)-Y^2(U-X)-XY^2\sin^2\Theta,\\ B&=U(U-X)^2-Y^2(U-X)-\tfrac{1}{2}XY^2\sin^2\Theta,\\ C&=(U-X)^3-Y^2(U-X).\end{aligned}\right\} \tag{4.66}$$

The solution of (4.62) has already been given at (4.51). The solution of (4.65) is

$$n^2=\frac{(U-X)(U^2-UX-Y^2)-\tfrac{1}{2}XY^2\sin^2\Theta\pm XY\{\tfrac{1}{4}Y^2\sin^4\Theta+\cos^2\Theta(U-X)^2\}^{\frac{1}{2}}}{U^2(U-X)-Y^2(U-X\cos^2\Theta)} \tag{4.67}$$

This can be shown to be the same as (4.47).

The form (4.47) of the solution is rather unusual first because of the 1 on the right-hand side and second because the square root is in the denominator. There is no simple corresponding form of the solution of (4.62) for the more general plasma (but see problem 4.13). This form seems to have appeared because it was derived from (4.12) without the need to invert the matrix, so that ρ appears in the denominator of (4.46).

Equation (4.65) can be treated as a quadratic for $1/n^2$ and solved to give

$$\begin{aligned} n^2&=\frac{2\varepsilon_1\varepsilon_2\varepsilon_3}{\varepsilon_1\varepsilon_2\sin^2\Theta+\tfrac{1}{2}\varepsilon_3(\varepsilon_1+\varepsilon_2)(1+\cos^2\Theta)\pm\{J^2\sin^4\Theta+\varepsilon_3^2(\varepsilon_1-\varepsilon_2)^2\cos^2\Theta\}^{\frac{1}{2}}}\\ &=\frac{(U-X)(U-X-Y)(U-X+Y)}{(U-X)(U^2-UX-Y^2)-\tfrac{1}{2}XY^2\sin^2\Theta\pm XY\{\tfrac{1}{4}Y^2\sin^4\Theta+(U-X)^2\cos^2\Theta\}^{\frac{1}{2}}}\end{aligned}$$

(electron plasma) (4.68)

Both (4.67) and (4.68) are useful. But (4.47) is much more widely used and is simpler for some forms of computation.

Another form of the dispersion relation is sometimes needed. To derive it we abandon the assumption that the wave normal is parallel to the z axis, and assume that it is in some general direction so that the factor $\exp(-\mathrm{i}knz)$ used in (4.1)–(4.7) is replaced by

$$\exp\{-\mathrm{i}k(xn_x+yn_y+zn_z)\}. \tag{4.69}$$

The refractive index is here treated as a vector $\boldsymbol{n}$ with components n_x, n_y, n_z; see § 5.1. Then if (4.69) is used in Maxwell's equations (2.44), with the constitutive relation (4.8), it gives

$$\boldsymbol{\Gamma}\boldsymbol{E}_0=\boldsymbol{\mathscr{H}}_0, \quad \boldsymbol{\Gamma}\boldsymbol{\mathscr{H}}_0=-\boldsymbol{\varepsilon}\boldsymbol{E}_0 \tag{4.70}$$

where $\mathbf{\Gamma}$ is the matrix

$$\mathbf{\Gamma} = \begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix}. \tag{4.71}$$

By multiplying (4.70) by $\mathbf{\Gamma}$ on the left, $\mathscr{H}_0$ may be eliminated to give

$$(\mathbf{\Gamma}^2 + \boldsymbol{\varepsilon})\boldsymbol{E}_0 = 0. \tag{4.72}$$

This can only have a solution if the determinant of the matrix on the left is zero; compare (4.60), (4.61). Hence

$$\begin{vmatrix} \varepsilon_{xx} - n_y^2 - n_z^2 & \varepsilon_{xy} + n_x n_y & \varepsilon_{xz} + n_x n_z \\ \varepsilon_{yx} + n_x n_y & \varepsilon_{yy} - n_x^2 - n_z^2 & \varepsilon_{yz} + n_y n_z \\ \varepsilon_{zx} + n_x n_z & \varepsilon_{zy} + n_y n_z & \varepsilon_{zz} - n_x^2 - n_y^2 \end{vmatrix} = 0. \tag{4.73}$$

This is the dispersion relation. The axes used here are different from those in (4.61) and so the elements of $\boldsymbol{\varepsilon}$ are different. (4.73) is a generalisation of (4.61) for the new axis system. It is used in § 14.2.

4.10. Zeros and infinity of refractive index. Equal refractive indices

The quadratic (4.65) with (4.66) or the solutions (4.68) show that one value of n^2, but only one, is zero when

$$X = U + Y, \quad \text{or} \quad X = U - Y, \quad \text{or} \quad X = U. \tag{4.74}$$

These are called 'cut-off' conditions, and the frequencies where they are satisfied are called cut-off frequencies. The corresponding conditions for the general plasma are, from (4.62)

$$\varepsilon_1 = 0 \quad \text{or} \quad \varepsilon_2 = 0 \quad \text{or} \quad \varepsilon_3 = 0 \tag{4.75}$$

When the third condition (4.74) is satisfied, the other non-zero value of of n^2 is $n^2 = 1$. For the general plasma when $\varepsilon_3 = 0$, the other non-zero n^2 is $2\varepsilon_1\varepsilon_2/(\varepsilon_1 + \varepsilon_2)$. An important feature of these conditions is that they are independent of the angle Θ between the wave normal and $\boldsymbol{Y}$.

If collisions are included, these conditions cannot be satisfied for any real values of X, but they are important when $X(h)$ and $Z(h)$ are treated as complex analytic functions of the height h regarded as a complex variable (in later chapters denoted by z). Points in the complex h plane where $n = 0$ are called reflection points.

The quadratic (4.65) with (4.66) or the solutions (4.67) show that one value of n^2, but only one, is infinite when

$$X = X_\infty \equiv U(U^2 - Y^2)/(U^2 - Y^2 \cos^2\Theta) \tag{4.76}$$

and for the general plasma

$$\tfrac{1}{2}(\varepsilon_1 + \varepsilon_2)\sin^2\Theta + \varepsilon_3 \cos^2\Theta = 0\,. \tag{4.77}$$

This condition is called 'resonance'. It does depend on the angle Θ. Its effect on wave propagation is discussed in §§ 18.13, 19.4–19.6.

The dispersive properties for radio waves, that is, the formulae for n^2 and ρ, are practically unaffected in most conditions when allowance is made for the electron and ion temperatures. But this is no longer true when the resonance condition is approached. Then (4.54) shows that for large n, E_z/E_x, E_z/E_y are of order n^2 and (4.2), (4.4) then show that $E_z/\mathscr{H}_x$, $E_z/\mathscr{H}_y$ are of order n. Thus the predominant field component is E_z and the wave resembles a longitudinal plasma wave. The full theory for a warm plasma (Ginzburg, 1970; Goland and Piliya, 1972) shows that the value of n does not actually become infinite when (4.76) or (4.77) is satisfied. It becomes very large and goes over continuously to the value for a plasma wave. See § 19.5 for further discussion.

Equation (4.46) shows that the two values of n^2 can be equal only when the two ρs are equal and the condition for this for an electron plasma was given at (4.26), (4.27). If collisions are neglected and X is real, this condition becomes

$$X = 1, \quad \sin\Theta = 0. \tag{4.78}$$

If $X(h)$ and $Z(h)$ are analytic functions of the height h (in later chapters denoted by z), the condition becomes

$$X(h) = 1 - \mathrm{i}\{Z(h) \pm Z_{\mathrm{t}}\} \tag{4.79}$$

where Z_{t} is given by (4.27). Points in the complex h plane where this is satisfied are called coupling points. They are studied in §§ 4.14, 16.5. There can be a coupling point on the real h axis only if

$$Z(h) = Z_{\mathrm{t}} \text{ where } X(h) = 1. \tag{4.80}$$

The condition for equal roots with (4.51), (4.62) for the more general plasma is discussed in § 13.9.

4.11. Dependence of refractive index on electron concentration 1. $Y < 1$

We shall now examine the dependence of the refractive index on electron concentration N for an electron plasma when collisions are neglected. This subject is particularly important in the study of the probing of the ionosphere by vertically incident radio waves, as in the ionosonde technique, § 1.7.

Since n^2 is real, it is useful to plot curves of n^2 versus X, which is proportional to N. Frequencies $f > f_H$ are considered first so that $Y < 1$. The form of the curves is then as shown in fig. 4.3. When $X = 0$ the medium is just a vacuum and $n^2 = 1$. Thus all the curves go through the point $X = 0$, $n^2 = 1$.

Consider now the special case $\Theta = \pm\pi/2$. Then (4.48) shows that the two solutions are

$$n_{\mathrm{O}}^2 = 1 - X, \quad n_{\mathrm{E}}^2 = 1 - \frac{X(1-X)}{1-X-Y^2}. \tag{4.81}$$

The subscripts are explained below. These two curves are shown as chain lines in fig. 4.3. The wave normal is perpendicular to Y and this case is therefore called 'transverse propagation'. It can be shown from (4.46), (4.29) that the two solutions (4.81) give

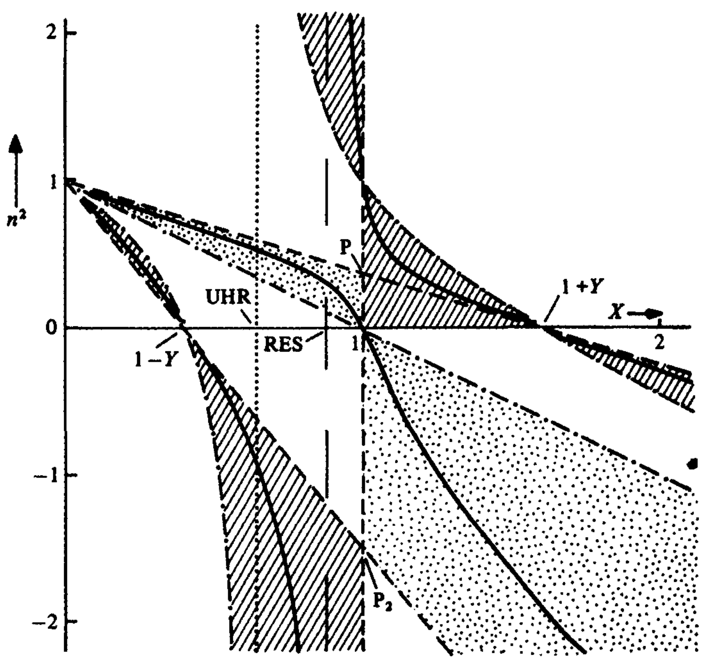
$$\rho_{\mathrm{O}} = 0, \quad \rho_{\mathrm{E}} = \infty \tag{4.82}$$

respectively. The first solution (4.81) is the same as for an isotropic plasma, and (4.82) with (4.53) (4.54) shows that $\boldsymbol{E}$ and $\boldsymbol{P}$ are parallel or antiparallel to $\boldsymbol{Y}$, that is to the earth's magnetic field. Thus the ordered part of the electron motions is unaffected by this field, and the wave is called the 'ordinary' wave. This is indicated by the subscript O.

The other wave is the 'extraordinary' wave and is indicated by the subscript E. It shows two zeros where $X = 1 \pm Y$ and a resonance where $X = 1 - Y^2$. This is known as the upper hybrid resonance; see § 3.11, item (6). The second equation (4.82) shows that $E_x = 0$, and (4.54) shows that

$$E_z/E_y = \pm \mathrm{i}XY/(1 - X - Y^2) \tag{4.83}$$

Fig. 4.3. Dependence of n^2 on X for a cold collisionless electron plasma when $Y<1$. In this example $Y \approx 0.6$. Broken lines are for $\sin\Theta = 0$, and chain lines for $\cos\Theta = 0$. The continuous lines are for an intermediate value of Θ, in this example 30°. Curves within the region shaded with dots are for the ordinary wave and curves in regions shaded with lines are for the extraordinary wave. RES is the value of X where $n^2 \to \infty$ for the extraordinary wave. UHR means 'upper hybrid resonance'.



(minus sign if $\sin\Theta$ is negative). Thus E_z and E_y are in quadrature.

When Θ differs slightly from $\pi/2$, the curve in fig. 4.3 for the ordinary wave moves away from the straight line $n^2 = 1 - X$, but it still goes through the zero at $X = 1$. Similarly the curves for the extraordinary waves go through the same zeros as before, but the resonance moves to a greater value of X, namely $(1 - Y^2)/(1 - Y^2\cos^2\Theta)$.

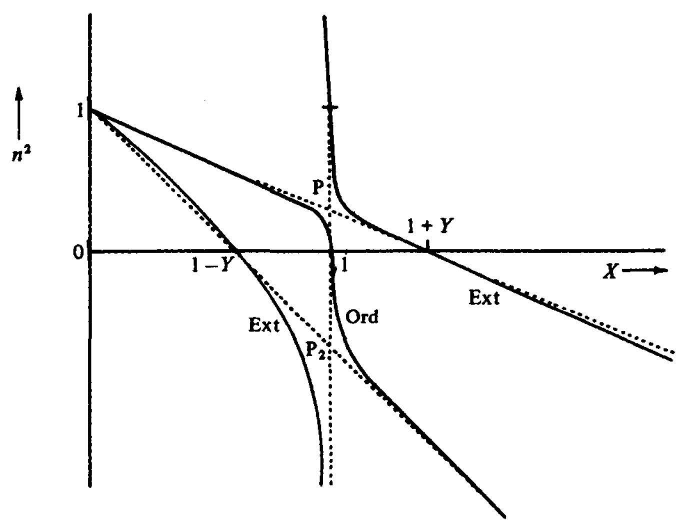
Next consider the special case $\sin\Theta = 0$. Then (4.48) shows that the two solutions are

$$n^2 = 1 - \frac{X}{1+Y}, \quad n^2 = 1 - \frac{X}{1-Y}\,. \tag{4.84}$$

These represent two straight lines through the zeros at $X = 1 \pm Y$, shown as broken lines in fig. 4.3. These solutions were discussed by Lorentz (1909, ch. IV). The wave normal is parallel or antiparallel to Y and this case is therefore called 'longitudinal propagation'. For upgoing waves near the north magnetic pole $\cos\Theta = +1$, and (4.46) shows that for the first of (4.84) $\rho = -\mathrm{i}$ and for the second $\rho = +\mathrm{i}$. These correspond to circularly polarised waves with a right-handed and a left-handed sense respectively. The formulae (4.84) do not show the zero at $X = 1$ nor the infinity, which (4.76) shows is also at $X = 1$. Neither do the two values of n^2 become equal where (4.79) is satisfied. This is because (4.84) and (4.78) were obtained from (4.48) and (4.20) respectively by setting $\sin^2\Theta = 0$, but this process is not valid when $X = 1$.

If $\sin\Theta$ is very small but not quite zero, the curves of n^2 vs X are as shown in fig. 4.4. The infinity for the extraordinary wave is close to $X = 1$, and both curves have sharp bends near $X = 1$. The curve for the ordinary wave may be followed from left to right. For $X < 1$, the corresponding value of ρ is close to $-\mathrm{i}$ but when the almost

Fig. 4.4. Dependence of n^2 on X when $\sin\Theta$ is very small, and $Y < 1$.



vertical part of the curve near $X = 1$ is traversed, ρ changes rapidly and goes over to a value close to $+\mathrm{i}$, which is retained for $X > 1$. It will be shown later (ch. 16) that this rapid change of polarisation is associated with strong coupling between the ordinary and extraordinary waves. If an ordinary wave enters a medium where it encounters an electron concentration that increases as it travels (increasing X) under the conditions of fig. 4.4, it would cause some extraordinary wave to be generated near $X = 1$. In the limit when $\sin\Theta = 0$, the extraordinary wave takes over completely where $X = 1$, and for $X > 1$ the only wave present would be the extraordinary wave.

This explains how the curves of fig. 4.4 go over, in the limit $\sin\Theta \to 0$, into the broken lines of fig. 4.3. The nearly vertical parts of the curves in fig. 4.4 become, in the limit, the vertical line $X = 1$ which is shown as a broken line in fig. 4.3 since it is really part of the curves for $\sin\Theta = 0$. On this line the refractive index n can take any value, and (4.52), (4.54) show that $E_x = E_y = P_x = P_y = 0$. The wave is a purely longitudinal wave and has all the properties of the plasma oscillation described near the end of § 3.8.

These results also show, by taking the curves of fig. 4.4 to the limit $\sin\Theta \to 0$, that in (4.84) when $X < 1$ the subscripts on n^2 should be $_{\mathrm{O},\,\mathrm{E}}$ respectively, and when $X > 1$ they should be $_{\mathrm{E},\,\mathrm{O}}$ respectively. The part of the vertical line $X = 1$ between $n^2 = -Y/(1-Y)$ at P_2 and $n^2 = Y/(1+Y)$ at P applies to the ordinary wave and the rest applies to the extraordinary wave. The curves for ordinary waves and extraordinary waves touch at P and P_2 where they have sharp bends and these two points are known as 'window points'. The point P is important in the theory of radio windows, §§ 17.6–17.9. The point P_2 is less important because there n^2 is negative and the waves are evanescent.

It can be shown, see § 5.2 item (12), that when $|\sin\Theta|$ lies between 0 and 1, the curves of n^2 vs X can never cross those for $\sin\Theta = 0$, or for $|\sin\Theta| = 1$. Thus they lie entirely within the regions that are shown shaded in fig. 4.3. For purely transverse propagation, $\cos\Theta = 0$, comparison of (4.81) and (4.47) shows that if the term $\pm\{\cdots\}^{\frac{1}{2}}$ in (4.48) is positive it gives the first equation (4.81), that is the value of n^2 for the ordinary wave. When Θ is now changed so that $\cos\Theta \neq 0$, the square root cannot change sign for any real $\cos\Theta$. Therefore, by continuity with the transverse case, we define n^2 for the ordinary wave as that value of (4.48) that has a positive value of the term $\pm\{\cdots\}^{\frac{1}{2}}$. Thus the curves for the ordinary wave in fig. 4.3 must lie in the region shaded with dots. Similarly n^2 for the extraordinary wave uses a negative value of the term $\pm\{\cdots\}^{\frac{1}{2}}$, and the curves must lie in the region shaded with lines. Further discussion of the terms 'ordinary' and 'extraordinary' is given in §§ 4.15, 4.16, 5.6.

The following further properties of the two refractive indices should be noted.
(a) When $X > 1$ the ordinary wave is evanescent for all Θ, since n^2 is negative.

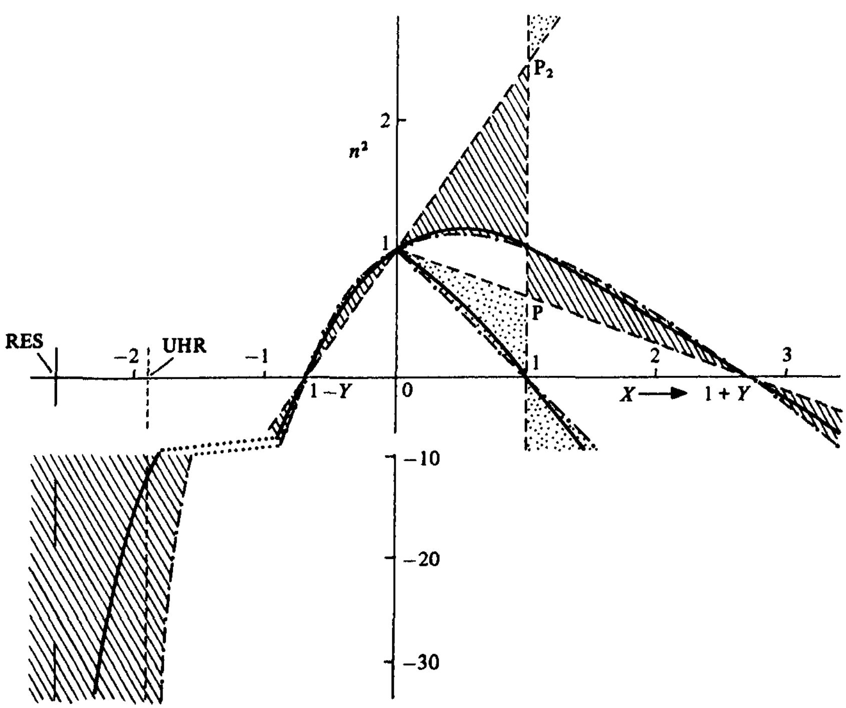
(b) The curve for the extraordinary wave has two branches. The wave for the right-hand branch in fig. 4.3 is called the Z-mode. The curve goes through the point $X = 1$, $n^2 = 1$ for all Θ. The wave is evanescent when $X > 1 + Y$.
(c) When $(1 - Y^2)/(1 - Y^2 \cos^2 \Theta) < X < 1$ the refractive index for the Z-mode is real and > 1 so that the wave velocity of the wave is less than the speed of light. Some authors call this wave the 'extraordinary slow wave' (Al'pert, 1983).

4.12. Dependence of refractive index on electron concentration 2. $Y > 1$

When the frequency is less than the electron gyro-frequency, about 1 MHz in the ionosphere at temperate latitudes, the form of the curves of n^2 vs X is somewhat different. At these low frequencies the effect of collisions becomes important but it is still of interest to consider first the consequences of neglecting collisions. The zeros and infinity of n^2 are still given by (4.74), (4.76) with $U = 1$ but the value of X at one of the zeros, $1 - Y$, is now negative. At the infinity X may be either positive in the range $X > 1$, or negative in the range $X < 1 - Y^2$. To illustrate these properties it is useful to draw the curves for negative as well as positive values of X, although negative values cannot occur at real heights. See figs. 4.5, 4.6.

The curves for purely transverse propagation are still given by (4.81) but now with

Fig. 4.5. Similar to fig. 4.3, but here $Y > 1$, and $|\cos \Theta| < 1/Y$. In this example $Y \approx 1.7$ and $\Theta \approx 72°$. See caption of fig. 4.3.



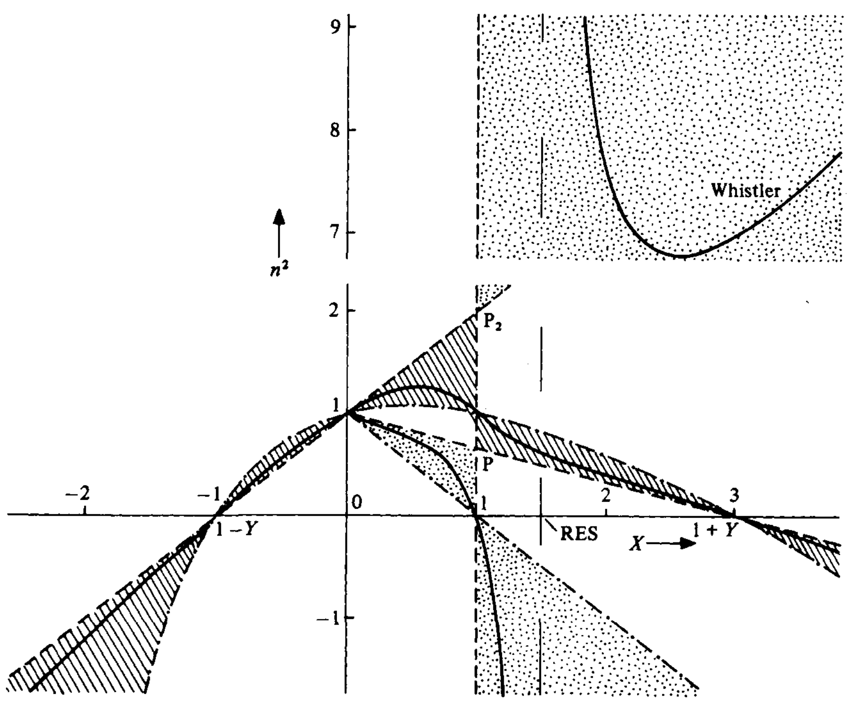
$Y > 1$. They are shown as chain lines in figs. 4.5, 4.6. Similarly the curves for purely longitudinal propagation are given by (4.84) together with the line $X = 1$; these are shown as broken curves in figs. 4.5, 4.6. On the line $X = 1$ there are two window points where $n^2 = Y/(Y-1)$ at P_2 and where $n^2 = Y/(Y+1)$ at P. For both of them n is real and they are associated with radio windows of two different types. See §§ 17.6–17.9.

The meaning of the terms 'ordinary' and 'extraordinary' is the same as for $Y < 1$ in § 4.11, so that n^2 for the ordinary wave uses a positive sign for the term $\pm\{\ldots\}^{\frac{1}{2}}$ in (4.47). Thus the curves for the ordinary wave lie in the region shaded with dots in figs. 4.5, 4.6, and the curves for the extraordinary wave in the region shaded with lines. The part of the line $X = 1$ between P and P_2 applies to the extraordinary wave, and the rest applies to the ordinary wave.

Fig. 4.5 illustrates the behaviour when $Y^2\cos^2\Theta < 1$, which includes the transverse and near transverse cases. The infinity, (4.76) with $U = 1$, is where X is negative and occurs for the extraordinary wave. For positive X each curve has only one branch. The ordinary wave is evanescent for all $X > 1$, and the extraordinary wave is evanescent for all $X > 1 + Y$.

Fig. 4.6 is for $Y^2\cos^2\Theta > 1$ which includes the longitudinal and near longitudinal

Fig. 4.6. Similar to fig. 4.3 but here $Y > 1$, and $|\cos\Theta| > 1/Y$. In this example $Y \approx 2.0$ and $\Theta \approx 30°$. See caption of fig. 4.3.



cases. The infinity is now where X is positive and > 1, and it occurs for the ordinary wave whose curve therefore has two branches. On the branch where $X > (Y^2 - 1)/(Y^2\cos^2\Theta - 1)$ the value of n^2 is positive and large, and tends to $+\infty$ when $X \to +\infty$. This is called the whistler branch and the wave is called the whistler mode. The phenomenon of whistlers is described in §13.8.

4.13. Effect of collisions included

When Z is not zero the two values of n^2 given by (4.47) are complex and we take $n = \mu - \mathrm{i}\chi$, (2.65), where μ and χ are real. It is now convenient to draw curves showing how μ and χ depend on X. These are slightly different in form from the curves of n^2 vs X used previously.

It was mentioned in §§2.13, and 4.2 that for a passive isotropic medium μ and χ must always be both positive (or both negative). This is still true when the earth's magnetic field is allowed for. It is necessary on physical grounds because the medium must always absorb energy from the wave, but it can be proved formally as follows. It is required to prove that $\mathrm{Im}(n^2)$ is negative, so that $\mathrm{Im}(1/n^2)$ is positive. The dispersion relation (4.62) may be written

$$\cos^2\Theta\left(\frac{1}{\varepsilon_1}-\frac{1}{n^2}\right)\left(\frac{1}{\varepsilon_2}-\frac{1}{n^2}\right)+\sin^2\Theta\left(\frac{1}{\varepsilon_3}-\frac{1}{n^2}\right)\left\{\frac{1}{2}\left(\frac{1}{\varepsilon_1}+\frac{1}{\varepsilon_2}\right)-\frac{1}{n^2}\right\}=0. \quad (4.85)$$

From the definitions (3.65)–(3.67) it can be shown that for real X_e, X_i, Z_e, Z_i each of $\varepsilon_1, \varepsilon_2, \varepsilon_3$ has a negative imaginary part and their reciprocals in (4.85) therefore have positive imaginary parts. If $1/n^2$ had a negative imaginary part, the imaginary part of the left side of (4.85) would be non-zero and positive, and the equation would not be satisfied. Hence $\mathrm{Im}(1/n^2)$ must be positive and non-zero if any collisions are included. An alternative proof that applies only to an electron plasma was given by Budden (1961a, §6.11).

4.14. The transition collision frequency

It was shown in §4.4 that for the electron plasma the two possible values of the polarisation ρ are both equal to ± 1 when $X = 1$, $Z = Z_t$ where $Z_t = \frac{1}{2}Y\sin^2\Theta/|\cos\Theta|$. This is simply the condition that the square root in the Appleton–Lassen formula (4.47) shall be zero, so that the two refractive indices are equal. Now it follows from (4.34), (4.35) and (4.38) that $\mathrm{Re}(\rho)$ and $\cos\Theta$ have opposite signs. Hence (4.46) shows that the one refractive index is then given by

$$n^2 = 1 - X/\{1 - \mathrm{i}(Z + Y|\cos\Theta|)\}. \quad (4.86)$$

For upgoing waves in the northern hemisphere $\cos\Theta$ is positive. For the transition condition, therefore, $\rho = -1$. This means that for both waves the electric vector is parallel to the plane $x = -y$.

It is shown later (§7.13) that the differential equations which govern the propagation of waves in a given direction (here taken to be parallel to the z axis) in a homogeneous medium are equivalent to a single differential equation of the fourth order. This should have four independent solutions, and it has been shown that in general this is so, for there are two solutions which represent progressive waves travelling in the direction of positive z, and two more for waves travelling in the opposite direction. For the transition condition described above, however, it might appear that there are only two independent solutions representing waves travelling in opposite directions, both with $\rho = -1$ and with n given by (4.86). But in this case two more solutions can be found, in which the electric vector has a component at right angles to the plane $x = -y$. This may be shown as follows.

Suppose that the medium is very nearly, but not quite, at the transition condition, so that $\rho = -1 \pm \varepsilon$ where ε is so small that its square and higher powers may be neglected, and let the associated values of the refractive index be n_+ and n_-. Then (4.46) gives

$$n_+^2 - n_-^2 = 2\mathrm{i}XY\varepsilon\cos\Theta\,(U - \mathrm{i}Y\cos\Theta)^{-2} \tag{4.87}$$

$$n_+ - n_- = \mathrm{i}XY\varepsilon\cos\Theta\, n^{-1}(U - \mathrm{i}Y\cos\Theta)^{-2} \tag{4.88}$$

where n (without subscript) is written for the average value of n_+ and n_-. Choose new axes x', y', z formed from the original axes (of § 4.3) by rotation through $-45°$ about the z axis, and let E'_x E'_y be the components of $\boldsymbol{E}$ parallel to the x' and y' axes, and let ρ' denote the new value of the polarisation ρ. Then (4.31) and (4.32) show that $\rho' = \pm\frac{1}{2}\varepsilon$. Let the two waves with these polarisations have equal amplitudes, so that their components E'_x annul each other when $z = 0$. Then

$$E'_x = A\{\exp(-\mathrm{i}kn_+z) - \exp(-\mathrm{i}kn_-z)\}, \tag{4.89}$$

where A is a constant. The other field components may be found from (4.16) and (4.2)–(4.5). To the first order in ε they are given by:

$$\left.\begin{aligned} E'_x &= -\mathrm{i}kz(n_+ - n_-)A\mathrm{e}^{-\mathrm{i}knz}, \\ E'_y &= A\varepsilon\mathrm{e}^{-\mathrm{i}knz}, \\ \mathscr{H}'_x &= -nA\varepsilon\mathrm{e}^{-\mathrm{i}knz}, \\ \mathscr{H}'_y &= A(n_+ - n_-)(1 - \mathrm{i}knz)\mathrm{e}^{-\mathrm{i}knz} \end{aligned}\right\} \tag{4.90}$$

Now let A tend to ∞, and ε tend to zero in such a way that $A\varepsilon$ remains constant and equal to B, say. Then by using (4.88) the field components may be written:

$$\left.\begin{aligned} E'_x &= BkXY\cos\Theta\, n^{-1}(U - \mathrm{i}Y\cos\Theta)^{-2}z\mathrm{e}^{-\mathrm{i}knz}, \\ E'_y &= B\mathrm{e}^{-\mathrm{i}knz}, \\ \mathscr{H}'_x &= -nB\mathrm{e}^{-\mathrm{i}knz}, \\ \mathscr{H}'_y &= \mathrm{i}BXY\cos\Theta\, n^{-1}(U - \mathrm{i}Y\cos\Theta)^{-2}(1 - \mathrm{i}knz)\mathrm{e}^{-\mathrm{i}knz} \end{aligned}\right\} \tag{4.91}$$

The wave field described by (4.91) has been derived by superimposing two progressive waves, travelling in the direction of positive z. The result is not a progressive wave since E'_x and $\mathscr{H}'_y$ have components with a factor z, but it is a solution of Maxwell's equations in this critical case. A fourth solution could be found that represented a similar wave travelling in the direction of negative z.

The wave (4.91) does not play any essential part in the theory of later chapters. The following is an example of a problem where it would be needed. Consider a homogeneous medium with a plane sharp boundary, and suppose that the transition condition holds for a direction normal to the boundary. A plane wave is incident normally on the boundary, and it is required to find the amplitude and polarisation of the reflected wave. The boundary conditions cannot in general be satisfied unless there are two waves in the medium, namely a linearly polarised progressive wave with refractive index given by (4.86), and a wave of the type (4.91). The detailed solution of this problem is complicated, however, and of no special practical interest.

4.15. The terms 'ordinary' and 'extraordinary'

When electron collisions are allowed for, it is desirable to have a formal definition of the terms 'ordinary' and 'extraordinary'. The following definition applies for a plasma in which the only effective charged particles are electrons, and only when the angle Θ between the wave normal and the earth's magnetic field is real. It still applies, however, when X and Z take complex values. This occurs when, in later chapters, the functions $X(h)$ and $Z(h)$ are continued analytically so as to use complex values of the height h, in later chapters denoted by z.

The Appleton–Lassen formula (4.47) and the formula (4.21) for ρ may be written

$$n^2 = 1 - \frac{X(U-X)}{U(U-X) - \frac{1}{2}Y^2\sin^2\Theta + S_{\mathrm{R}}}, \quad \rho = \frac{\mathrm{i}(\frac{1}{2}Y^2\sin^2\Theta - S_{\mathrm{R}})}{Y(U-X)\cos\Theta} \tag{4.92}$$

where

$$S_{\mathrm{R}}^2 = \tfrac{1}{4}Y^4\sin^4\Theta + Y^2(U-X)^2\cos^2\Theta. \tag{4.93}$$

Definition: When $\mathrm{Re}(S_{\mathrm{R}})$ is positive, the values of n and ρ from (4.92) are written n_{O}, ρ_{O} and are the refractive index and polarisation of the ordinary wave. When $\mathrm{Re}(S_R)$ is negative, the values are written n_{E}, ρ_{E} and are the refractive index and polarisation of the extraordinary wave.

When $\cos\Theta \to 0$, that is transverse propagation, this definition gives $n_{\mathrm{O}}^2 = 1 - X/U$, $\rho_{\mathrm{O}} = 0$, so it accords with the definition given in §4.11.

When Θ, X, Z are all real, and $Z < Z_t$, (4.26)–(4.28), and $\sin\Theta \neq 0$, the real part of S_{R}^2 is positive and never zero. Then S_{R} necessarily has a non-zero real part, and there is no ambiguity in the terms ordinary and extraordinary. When $Z > Z_t$, or when $\sin\Theta = 0$ and $Z = 0$, there is ambiguity when $X \equiv 1$. This is discussed in the following section. When X and Z are complex, $\mathrm{Re}(S_{\mathrm{R}})$ can sometimes be zero and there is then ambiguity in the definition. When $X(z)$ and $Z(z)$ are functions of the height z as in later chapters, the condition $\mathrm{Re}(S_R) = 0$ is the equation of a line in the complex z plane. It is sometimes convenient to use this line as a branch cut; see §§ 16.5–16.7.

The above definition is useful for radio waves vertically incident on a horizontally stratified ionosphere, for then Θ has a fixed real value. For oblique incidence the problem is more complicated. Instead of the refractive indices $\pm n_O$, $\pm n_E$, the four roots q of the Booker quartic equation must be used; ch. 6. The angle Θ between the wave normal and the earth's magnetic field now has different values for the four waves. It is in general complex and varies with height. The above definition could in principle still be used but is rarely or never invoked, because the need for a precise definition does not arise.

At extremely low frequencies where the effect of heavy ions is important a definition similar to that given above can be formulated; see, for example, Al'pert, Budden *et al.* (1983, § 8(a)). But at these low frequencies the waves are usually described by other names (Booker, 1984), and the terms ordinary and extraordinary are not often used.

4.16. Dependence of refractive index on electron concentration 3. Collisions allowed for

To extend the results shown in figs. 4.3–4.6 so as to allow for electron collisions, it would be necessary to plot curves of both μ and χ (2.65) versus X. To cover all cases of interest a very large number of curves would be needed and no attempt has been made to supply them in this book. Some curves of this kind have been given by Booker (1934), Ratcliffe (1959), Budden (1961a), Ginzburg (1970), Kelso (1964), Rawer and Suchy (1967). When a computer is available, μ and χ can be calculated very quickly and it is easy to plot the curves as needed.

For purely transverse propagation the first formula (4.81) is for the ordinary wave and is now replaced by (4.9). This is the same as for any wave in an isotropic electron plasma, and typical curves of μ and χ for this case are shown in fig. 4.1. The second formula (4.81) is for the extraordinary wave and is now replaced by

$$n_E^2 = 1 - \frac{X(U - X)}{U(U - X) - Y^2}. \tag{4.94}$$

Curves of μ and χ were given by Budden (1961a) for the two typical cases $Y < 1$ and $Y > 1$.

For purely longitudinal propagation the formulae (4.84) are replaced by

$$n^2 = 1 - \frac{X}{U + Y}, \quad n^2 = 1 - \frac{X}{U - Y} \tag{4.95}$$

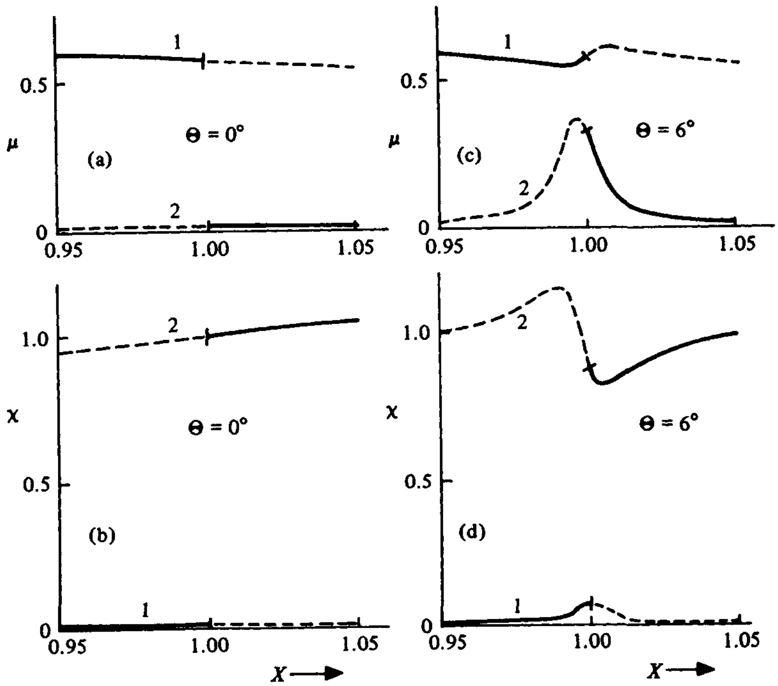
The dependence on X of μ and χ from the first formula (4.95) is given by fig. 4.1 provided that the abscissa X is replaced by $X/(1 + Y)$ and the numbers by the curves are the values of $Z/(1+Y)$. The same applies to the second formula (4.95) if $Y<1$; the abscissa is now $X/(1-Y)$ and the numbers by the curves are the values of $Z/(1-Y)$. If $Y>1$ a new set of curves is needed, as given by Budden (1961a).

It was shown in § 4.11 that when $Z=0$ and $\sin\Theta$ is zero, the curve of n^2 versus X makes a steep traverse near $X=1$ from the curve given by the first formula (4.84) to that given by the second, and the two sections of the resulting curve are joined by a segment of the vertical line $X=1$. When $\sin\Theta$ is small but not quite zero this steep traverse still occurs as shown in fig. 4.4 where the curve is nearly vertical near $X=1$. When $Z\neq 0$ and when $\sin\Theta$ is zero or small, this steep traverse does not occur. It appears only when Θ exceeds a transition value Θ_t given by

$$\sin^2\Theta_t/|\cos\Theta_t| = 2Z/Y. \qquad (4.96)$$

This is the same as (4.27) except that the subscript t appears on Θ, and not on Z. In deriving (4.27), Θ was regarded as held constant and a transition value of Z was sought. Here we suppose that Z is held constant and seek a transition value Θ_t of Θ. When (4.27) or (4.96) is satisfied the two values of the polarisation ρ are both equal to -1 (for upgoing waves in the northern hemisphere) and the two refractive indices are equal.

Fig. 4.7. Shows how μ and χ depend on X for a cold electron plasma with $Y=0.5$, $Z=0.007\,655$, $\Theta_t=10°$. These curves should be studied as part of the sequence, figs. 4.7–4.11. In (a), (b) $\Theta=0°$. In (c) (d) $\Theta=6°$.



Figs. 4.7–4.11 show curves of μ and χ versus X for a fixed value of Z and for a sequence of values of Θ. They cover a small range of X near $X = 1$ where the transition occurs. In this example the transition value of Θ is $\Theta_t = 10°$. Figs. 4.7 (a, b) are for $\Theta = 0$. The values of μ and χ are very close to those that would be obtained from the straight lines (4.95) in fig. 4.3, where U was equal to 1. Thus for the first expression (4.95) with $+ Y$, upper straight broken line in fig. 4.3, n^2 is almost real and positive but now the non-zero Z gives it a small imaginary part, so χ is non-zero but small. The resulting curves are labelled 1 in figs. 4.7(a, b). For the second expression (4.95) with $- Y$, lower straight broken line in fig. 4.3, n^2 is almost real and negative, but again it now has a small imaginary part. Thus μ is non-zero but small. The curves in figs. 4.7(a, b) are labelled 2. There is nothing in these two figures to correspond to

Fig. 4.8. Dependence of μ and χ on X for the same plasma as in fig. 4.7. The right-hand figure shows the complex ρ plane with the values of ρ for the two waves. The arrows show the direction of increasing X and the bars are where $X = 1$. Here $\Theta = 9°$.

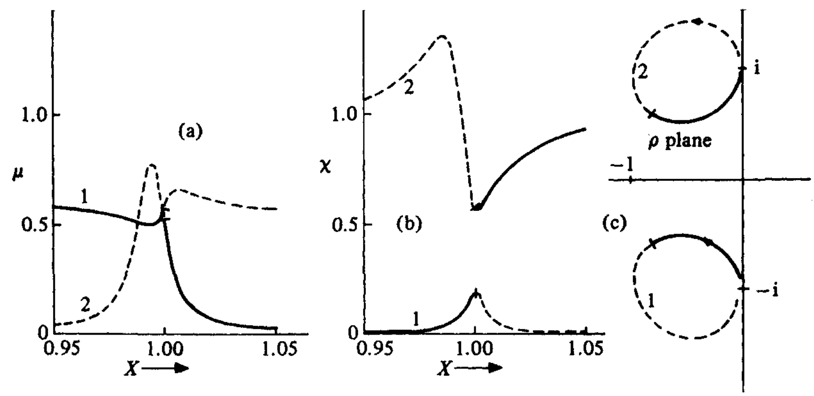
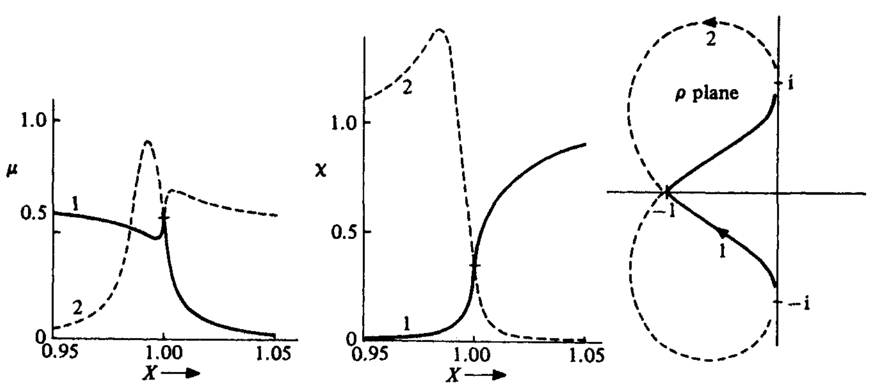


Fig. 4.9. Similar to fig. 4.8 but for the transition case $\Theta = \Theta_t = 10°$.



the vertical line $X = 1$ in fig. 4.3. For curves 1 the polarisation ρ is $-\mathrm{i}$ for all X, and for curves 2 it is $+\mathrm{i}$.

The wave corresponding to the first expression (4.95), described by curves 1, was, in § 4.11, called the 'ordinary wave' for $X < 1$, and the 'extraordinary wave' for $X > 1$. A wave with $\rho = -\mathrm{i}$ which enters a plasma where $X < 1$, with $\Theta = 0$, is an ordinary wave. Suppose now that the value of X increases with distance as the wave travels. The values of μ and χ are given by curves 1 of figs. 4.7(a, b), and they are smoothly varying functions of distance. If the medium varies very slowly with distance, the wave travels on without reflection, and it continues to do so when it enters a region where $X > 1$. After it has passed the point where $X = 1$, however, it is called an 'extraordinary wave'. There is no physical process associated with this change, and in particular no 'mode conversion', and no change of polarisation. All that has changed is the adjective used to describe the wave. See § 16.6 for an example.

Figs. 4.7(c, d) show similar curves of μ and χ versus X for $\Theta = 6°$. This again is less than Θ_t. The curves are continuous and the same remarks apply. The curves for the

Fig. 4.10. Similar to fig. 4.8 but here $\Theta = 10.5°$, slightly greater than the transition value $\Theta_t = 10°$.

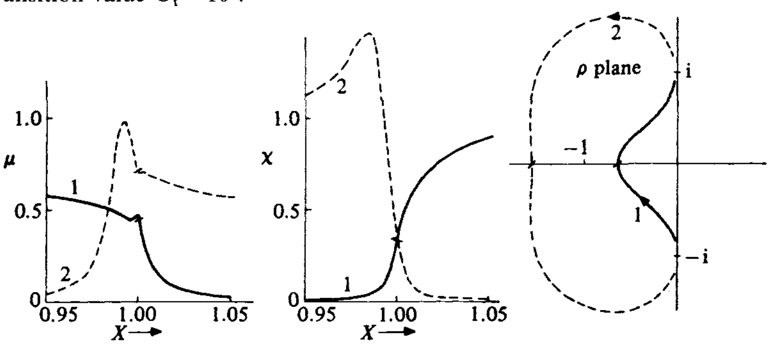
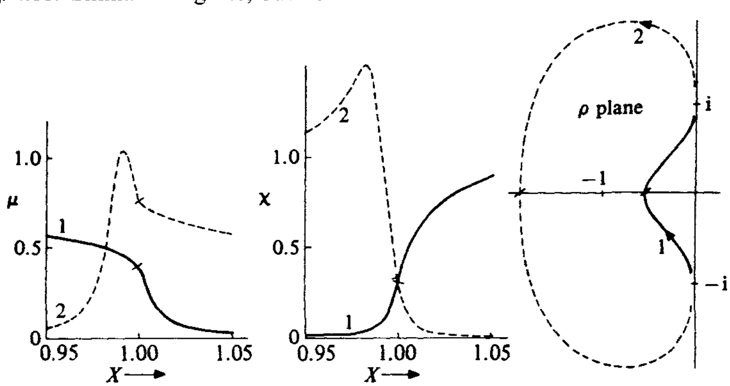


Fig. 4.11. Similar to fig. 4.8, but here $\Theta = 11°$.



ordinary wave are shown as continuous lines and those for the extraordinary wave as broken lines, and this convention applies for the whole sequence, figs. 4.7–4.11. There is a change of description in figs. 4.7(c, d) when X passes the value 1, as before. The polarisations ρ remain close to $-\mathrm{i}$ and $+\mathrm{i}$ for curves 1 and 2 respectively. There are now small departures from these values when X is near to 1.

Fig. 4.8 shows curves for $\Theta = 9°$ which is very slightly less than Θ_t. The general behaviour is still similar to that of the cases in fig. 4.7. The two curves for μ in fig. 4.8(a) now cross each other twice. It is the crossing nearest to $X = 1$ that is important in the present discussion. The curves for χ do not cross. Fig. 4.8(c) shows how the two polarisations ρ change when X changes. When X is near 1 they now depart appreciably from the values $\pm\mathrm{i}$ and have both moved nearer to the value $\rho = -1$. For figs. 4.7, 4.8 the conditions are sometimes said to be 'less than critical'. This term is explained in §§16.5, 16.6.

Fig. 4.9 shows curves for the transition case $\Theta = \Theta_t = 10°$. Now the curves for μ have a crossing where $X = 1$, and those for χ have now touched each other where $X = 1$. The curves for ρ both go through the value $\rho = -1$ when $X = 1$. Fig. 4.10 shows what happens when $\Theta = 10.5°$, slightly greater than Θ_t. The curves have separated again but are now linked in a different way. The curves for χ now cross near $X = 1$, and those for μ do not cross there. Both curves are continuous at $X = 1$. The wave that is ordinary for $X < 1$ now remains an ordinary wave for $X > 1$, and similarly for the extraordinary wave. There is now no change of the descriptive adjectives. Fig. 4.10 shows that the polarisation of the ordinary wave now goes over from near $-\mathrm{i}$ when $X < 1$, to near $+\mathrm{i}$ when $X > 1$, and the opposite transition occurs for the extraordinary wave. The functions μ and χ change rapidly as X varies, near $X = 1$, and this is the counterpart of the steep traverse near $X = 1$ in fig. 4.4. Finally fig. 4.11 shows the behaviour for $\Theta = 11°$ where the condition $\Theta > \Theta_t$ is well established. The behaviour of μ, χ and ρ is now just what might be expected if the continuous curves in fig. 4.3 near $X = 1$ were used to plot μ and χ, but modified to include a non-zero value of Z. For figs. 4.10, 4.11 the conditions are sometimes said to be 'greater than critical', and this is explained in §§16.5, 16.6.

The transition phenomena just described, and illustrated by figs. 4.7–4.11 were for the special case $Y = \frac{1}{2}$ and are typical for $Y < 1$. When $Y > 1$ the phenomena are very similar but there are minor differences. For example when Θ is slightly less than Θ_t, the two curves $\chi(X)$ cross near $X = 1$ and the two curves $\mu(X)$ do not cross. When $\Theta > \Theta_t$ the $\mu(X)$ curves cross and the $\chi(X)$ curves do not cross.

For a plasma with the effect of ions allowed for, this type of transition is important in the theory of ion-cyclotron whistlers §13.9.

The interchange of the adjectives 'ordinary' and 'extraordinary' at $X = 1$ for $\Theta < \Theta_t$ could be avoided by adopting different definitions of these adjectives when $X > 1$. But this does not remove the need for an interchange. It would simply mean that if X is held constant at some value > 1, and if Θ is then varied continuously from $< \Theta_t$ to $> \Theta_t$, the two adjectives would have to be interchanged when Θ passes the value Θ_t. There can be no choice of definitions that will altogether avoid the need for aninterchange. The choice made here is the one adopted in figs. 4.7–4.11. It is used throughout this book and by most authors, but not all.

The transition just described occurs when two solutions of the quadratic (4.20) for ρ are equal, and the condition is given by (4.25) which with (4.27) leads to

$$X = 1 - \mathrm{i}(Z \pm Z_t). \tag{4.97}$$

This represents two points in the complex X plane. The one with the + sign can never be real, but the other is real when $Z = Z_t$, which is equivalent to (4.96), and this is the transition condition. The point $X = 1 - \mathrm{i}(Z - Z_t)$ is a branch point of the function $n(X)$. It is in the negative imaginary half of the complex X plane when $Z > Z_t$ or equivalently $\Theta < \Theta_t$. As Z is decreased, or Θ increased, so that the transition condition is approached, this branch point moves up to the real X axis and lies on it at $X = 1$ at the exact transition. When $Z < Z_t$, or equivalently $\Theta > \Theta_t$, it is in the positive imaginary half of the complex X plane. This branch point is called a 'coupling point'. In a medium which is not very slowly varying near $X = 1$, one of the two waves can give rise to some of the other wave by a mode conversion process called coupling. It is studied in §§ 16.5–16.7.

In this section we have studied the behaviour of $\mu(X)$ and $\chi(X)$ when the point X moves along the line $\mathrm{Im}(X) = 0$. The two values of $n = \mu - \mathrm{i}\chi$ are those of the double valued complex function $n(X)$. If the branch point is where $\mathrm{Im}(X)$ is negative, the two curves for $\mu(X)$ cross each other, near $X = 1$, and the curves for $\chi(X)$ do not cross. When the branch point is on the line $\mathrm{Im}(X) = 0$, both curves cross at $X = 1$ where the two complex values of n are equal. When the branch point is where $\mathrm{Im}(X)$ is positive, the curves for $\chi(X)$ cross and those for $\mu(X)$ do not cross near $X = 1$. It can be shown that this behaviour always occurs for any two-valued complex function with a single branch point. It was studied for magnetoionic theory by Booker (1934).

For any X and Z, including real values, the two refractive indices are equal when Θ has some complex value, Θ_g say. If curves of μ and χ are plotted against $\mathrm{Im}(\Theta)$, with $\mathrm{Re}(\Theta)$ held constant, their behaviour when $\mathrm{Im}(\Theta)$ is near to $\mathrm{Im}(\Theta_g)$ would be similar to that of the curves $\mu(X)$ and $\chi(X)$ described above. The curves, for μ or χ, that cross when $\mathrm{Re}(\Theta) < \mathrm{Re}(\Theta_g)$ would not cross when $\mathrm{Re}(\Theta) > \mathrm{Re}(\Theta_g)$ and vice versa. There could again be a need to interchange the adjectives 'ordinary' and 'extraordinary' when $\mathrm{Im}(\Theta) > \mathrm{Im}(\Theta_g)$.

4.17. Approximations for refractive indices and wave polarisations

Various approximations have been used for n^2 and for ρ. These were useful in the days before computers and calculators were available, when calculations with the exact formulae were laborious. In theoretical physics many of the approximations of a function, $f(z)$ say, for values of z near to some specified value z_0, use expansions in powers of $z - z_0$. One type is an asymptotic approximation. This is discussed in §§ 8.11, 8.17 and is not considered further in the present section. Another type applies when $f(z)$ is analytic for z equal to or near to z_0. Then $f(z)$ is expanded in a Taylor series whose successive terms are $f(z_0)$, $(z - z_0)(\mathrm{d}f/\mathrm{d}z)_0$, $\frac{1}{2}(z - z_0)^2(\mathrm{d}^2f/\mathrm{d}z^2)_0$, etc., where the subscript 0 indicates that the value at $z = z_0$ is used. The approximation consists in retaining only two or three terms of the series. If r terms are retained, the approximation gives $f(z_0)$ and its derivatives up to $(\mathrm{d}^{r-1}f/\mathrm{d}z^{r-1})_0$ exactly. When $z \neq z_0$ the error in the approximation is of order $(z - z_0)^r$ and therefore small when $|z - z_0|$ is small. Approximations that are expansions with this analytic property are widely used in theoretical physics and there are many examples in this book.

Any approximation, whether based on an expansion or on some other method, cannot be correct under all conditions. The most important thing to know about it is where it fails, so that errors are avoided.

In magnetoionic theory one group of approximate formulae are called quasi-longitudinal (QL), and quasi-transverse (QT) approximations. They have been known for some time and are here called the 'conventional' QL and QT approximations. They are used only for an electron plasma, and are as follows:
QL:

$$n^2 \approx 1 - X/(U \pm Y\cos\Theta), \quad \rho \approx \mp \mathrm{i}. \tag{4.98}$$

If $(1 - X)(1 - Y)\cos\Theta$ is positive, the upper signs are for the ordinary wave and the lower for the extraordinary wave. If it is negative, the upper signs are for the extraordinary and the lower for the ordinary wave.
QT, ordinary wave:

$$n^2 \approx 1 - X/\{U + (U - X)\cot^2\Theta\}, \quad \rho \approx 0 \tag{4.99}$$

QT, extraordinary wave:

$$n^2 \approx 1 - X(U - X)/\{U(U - X) - Y^2\sin^2\Theta\}, \quad 1/\rho \approx 0. \tag{4.100}$$

The QL approximations (4.98) are exactly correct when $\sin\Theta = 0$, and the QT approximations (4.99), (4.100) are exactly correct when $\cos\Theta = 0$. But they fail in the following conditions:

(a) Except for (4.99), they do not make n^2 zero at the correct values of X, and the values of X where $n^2 = 0$ depend on Θ.

(b) They do not make n^2 infinite at the correct value of X.

(c) They do not give the correct derivatives $\mathrm{d}(n^2)/\mathrm{d}(\sin^2\Theta) = n\mathrm{d}^2n/\mathrm{d}\Theta^2$ at $\sin\Theta = 0$ for QL, nor $\mathrm{d}(n^2)/\mathrm{d}(\cos^2\Theta) = n\mathrm{d}^2n/\mathrm{d}\Theta^2$ at $\cos\Theta = 0$ for QT. Thus they do not give the correct values of the curvature of the refractive index surface in these two cases.

It might be thought that the term 'quasi-longitudinal' implies that the approximation is good in conditions at and near to longitudinal, that is near $\sin^2\Theta = 0$, and that it gives correct values of one or two of the derivatives when $\Theta \to 0$. The term 'quasi-transverse' could be taken to have similar implications. But the failure (c) shows that this is not so. In the analytic sense mentioned above, the conventional QL and QT approximations are not correct. In spite of their failures they have been used successfully but it seems that no clear rule has ever been given to show when they are useful and when they fail.

The form (4.63) of the dispersion relation shows that near $\sin^2\Theta = 0$, n^2 is an analytic function of $\sin^2\Theta$ so that it can be expanded in a series of powers of $\sin^2\Theta$. Similarly (4.64) shows that near $\cos^2\Theta = 0$, n^2 is an analytic function of $\cos^2\Theta$ and a similar expansion is possible. These expansions were given by Budden (1983), both for an electron plasma and for the more general plasma. They will be referred to here by the abbreviations QL2 and QT2. They are as follows:

QL2:

$$n^2 \approx \left(1 - \frac{X}{1+Y}\right)\left\{1 - \frac{XY\sin^2\Theta}{2(U-X)(U+Y)} + \dots\right\} \tag{4.101}$$

$$n^2 \approx \varepsilon_1\left\{1 - \frac{\varepsilon_1 - \varepsilon_3}{2\varepsilon_3}\sin^2\Theta + \dots\right\}. \tag{4.102}$$

These are for one of the two characteristic waves. For the other, the sign of Y is reversed in (4.101), and ε_1 is replaced by ε_2 in (4.102).

QT2 ordinary wave:

$$n^2 \approx \left(1 - \frac{X}{U}\right)\left\{1 + \frac{X}{U}\cos^2\Theta + \dots\right\} \tag{4.103}$$

$$n^2 \approx \varepsilon_3\left\{1 - \frac{(\varepsilon_1 - \varepsilon_3)(\varepsilon_2 - \varepsilon_3)}{J}\cos^2\Theta + \dots\right\}. \tag{4.104}$$

QT2 extraordinary wave:

$$n^2 \approx \frac{(U-X)^2 - Y^2}{U(U-X) - Y^2}\left\{1 - \frac{X(U-X)}{U(U-X) - Y^2}\cos^2\Theta + \dots\right\} \tag{4.105}$$

$$n^2 \approx \frac{2\varepsilon_1\varepsilon_2}{\varepsilon_1 + \varepsilon_2}\left\{1 - \frac{\varepsilon_3(\varepsilon_1 - \varepsilon_2)^2}{2(\varepsilon_1 + \varepsilon_2)J}\cos^2\Theta + \dots\right\}. \tag{4.106}$$

The coefficients of the terms in $\sin^4\Theta$ and $\cos^4\Theta$ in these expansions were given by Budden (1983), who also gave the corresponding expansions for the wave polarisations ρ, but these details are not needed here.

The formulae (4.101)–(4.106) give the zeros of n^2 at the correct values of X and they give the correct values of $n\,\mathrm{d}^2n/\mathrm{d}\Theta^2$ for $\sin\Theta = 0$ (QL2) and $\cos\Theta = 0$ (QT2) respectively. Thus they do not fail under conditions (a), (c) above. But they do not make n^2 infinite at the correct value of X and so they fail under condition (b).

The radius of convergence of the QL2 expansion is the smallest value of $|\sin^2\Theta|$ for which n^2 has a singularity, and there are two of these. First there is a pole where the denominator of (4.67) or of (4.51) is zero, but this appears in only one of the two solutions, namely the extraordinary wave if $Y < 1$, and the ordinary wave if $Y > 1$. It gives as a convergence criterion:

For QL2:

$$|\sin^2\Theta| < \left|\frac{(U-X)(U^2-Y^2)}{XY^2}\right|, \quad |\sin^2\Theta| < \left|\frac{\varepsilon_3}{G}\right| \tag{4.107}$$

The second singularity is the branch point where the square root in (4.47) or (4.51) is zero. This gives as a convergence criterion:

For QL2:

$$\begin{aligned} |\sin^2\Theta| &< 2\left|\frac{U-X}{Y}[U-X\pm\{(U-X)^2-Y^2\}^{\frac{1}{2}}]\right| \\ |\sin^2\Theta| &< \left|\frac{\varepsilon_3(\varepsilon_1-\varepsilon_2)^2}{2J}[\varepsilon_3\pm\{\varepsilon_3^2-4J^2(\varepsilon_1-\varepsilon_2)^{-2}\}^{\frac{1}{2}}]\right| \end{aligned} \tag{4.108}$$

where it is implied that the smaller values are used.

In a similar way we obtain convergence criteria for $|\cos^2\Theta|$ in the QT2 approximations. The radius of convergence set by the pole, for one wave only, is:

For QT2:

$$\begin{aligned} |\cos^2\Theta| &< \left|\frac{U(U^2-UX-XY^2)}{XY^2}\right| \\ |\cos^2\Theta| &< \left|1+\frac{\varepsilon_3}{G}\right|. \end{aligned} \tag{4.109}$$

The range given by the branch point is a rather complicated expression not needed here but given by Budden (1983).

The convergence criteria give some indication of where the QL2 and QT2 approximations may and may not be used. Budden (1983) gave examples to illustrate that for frequencies in the radio range, that is not much less than half the electron gyro-frequency, the QL2 and QT2 versions are considerably better than the conventional QL and QT. Johler and Walters (1960) studied the conventional QL approximation for very low frequencies in the range 10 to 80 kHz. They describe it as crude and warn that it should be used only with considerable caution.

Consider now the application of these formulae to extremely low frequencies, in the range studied in magnetohydrodynamics. For frequencies less than the smallest ion gyro-frequency ε_3 is very large and negative, and ε_1, ε_2 are positive and approximately equal (see fig. 3.2). Now (4.109) second equation gives an extremely small convergence range for QT2, for the wave that has an infinity of n^2 (Booker and Vats, 1985). This is a clear example where QT2 is not useful. But for the other wave there is no infinity of n^2 and (4.109) does not apply to it. For this wave QT2 is still useful.

Other forms of approximation are possible that avoid failure at an infinity of n^2. Instead of finding an approximation for n^2, the method is to find an approximation for $n^2\{\frac{1}{2}(\varepsilon_1+\varepsilon_2)\sin^2\Theta+\varepsilon_3\cos^2\Theta\}$. Here it is the second factor, in brackets, that is zero when one n^2 is infinite. This method was used by Booker and Vats (1985). The author is greatly indebted to Professor H.G. Booker for supplying a copy of this paper in advance of publication, and for a stimulating correspondence. Booker and Vats showed that, for the wave that has an infinity, high accuracy is attained by using either the QL or QT form of this version, over a wide range of frequency and angle Θ. But for the other wave their method gives infinite n^2 where in fact there is no infinity.

Another form of approximation should be mentioned. For an electron plasma the dispersion relation is (4.65), (4.66), and it is easy to show that as $\omega\to 0$, $X\to\infty$, both values of n^2 tend to infinity. But (4.65) can be written as an equation for n^2/X, and its solutions expanded in powers of $1/X$. For a collisionless plasma, $U=1$, the first terms of this expansion for the two waves are

$$n^2=-X/(1\pm Y\cos\Theta). \tag{4.110}$$

The radius of convergence of the expansion imposes the restriction that

$$1/X<|(Y^2\cos^2\Theta-1)/(Y^2-1)|. \tag{4.111}$$

The formula (4.110), with the lower sign, has been found useful for studying the propagation of whistlers (§ 13.8) in the magnetosphere. Here Y is of order 10 to 100 and X is of order 10^2 to 10^4. The condition (4.111) is amply satisfied except when $\cos\Theta$ is very close to $1/Y$. It can be shown that the second term of the expansion is small. The value $\cos\Theta=1/Y$ makes n^2 infinite in (4.110). The correct value is $(1/Y)\{1+(Y^2-1)/X\}^{\frac{1}{2}}$ so the error is small when X is large.

The formula (4.110) happens to be the same as the conventional QL approximation, (4.98) without the 1, and some authors have regarded it as an angular approximation rather than as depending on a large value of X. Its simplicity makes it especially useful. For example the formulae (5.75) for the group refractive index, and (5.83) for the angle α between the ray and the wave normal are simple, and Gendrin (1960, 1961) has used them in a study of whistler propagation. Booker (1984) has shown that the formula $\mathscr{U}=\partial\omega/\partial\boldsymbol{\kappa}$, (5.72), in this case can be expressed very simply in Cartesian coordinates, whereas it is not so easy with the exact formula; see § 5.8.

The formula (4.110) may be unreliable if the frequency is too great, so that X is not large or if the frequency is so small that the effect of heavy ions cannot be ignored. The useful frequency range is roughly from the lower hybrid resonance frequency (3.74) to the electron gyro-frequency.

Approximations for n^2 are sometimes needed for very high frequencies so that X, Y and Z are very small. For small X both values of n^2-1 are of order X. Suppose that X^2 and higher powers can be neglected. Then the dispersion relation (4.65)

(4.66) is expressed as a quadratic for $(n^2-1)/X$ and in the coefficients we set $X=0$. The solution is

$$n^2 - 1 \approx X\frac{\frac{1}{2}Y^2\sin^2\Theta - U^2 \pm Y\{\frac{1}{4}Y^2\sin^4\Theta + U^2\cos^2\Theta\}^{\frac{1}{2}}}{U(U^2-Y^2)}. \tag{4.112}$$

This result is used in the study of limiting polarisation, § 17.11. At high frequencies, Z and Y are small. If $|\cos\Theta/\sin^2\Theta| \gg \frac{1}{2}Y$, and if Z and Y^2 and higher powers of Y can be neglected, this gives

$$n^2 \approx 1 - X(1 \mp Y\cos\Theta); \quad n \approx 1 - \tfrac{1}{2}X(1 \mp Y\cos\Theta). \tag{4.113}$$

The upper sign is for the ordinary wave. This result is used in the study of Faraday rotation, § 13.7.

This discussion has shown that there are many ways of finding approximations for n^2 and ρ, and some of them are useful for specified applications. The subject has been fully discussed by Booker (1984). In nearly all cases it is advisable, before using them, to make some check of accuracy and this usually means comparing them with the exact formula. But then the exact formula might just as well be used from the start and there is no need to go to the trouble of making a check. Even the most complicated forms of the cold plasma dispersion relation can be solved very quickly with modern computers. In the author's opinion approximations of the kind described in this section are rarely needed for frequencies in the range used for radio communication, that is when X and Y are neither very large nor very small, and are likely to go out of use. For warm and hot plasmas the situation is quite different. The full form of the dispersion relation is extremely complicated and approximations are invariably used.

PROBLEMS 4

4.1. An elliptically polarised plane electromagnetic wave has its wave normal parallel to the z axis. Its wave polarisation is $E_y/E_x = R$ (complex). What are the axis ratio and the square of the length of the semi-major-axis of the polarisation ellipse, in terms of $|E_x|$ and R?

4.2. The wave of problem 4.1 is incident normally on a cold magnetoplasma in which the characteristic polarisations are ρ and $1/\rho$. Find the ratio of the major axes of the polarisation ellipses of the two resulting waves in the plasma, and check that the answer is right for the special cases $R=\rho$ or $1/\rho$.

4.3. For a plane progressive radio wave in a cold collisionless electron magnetoplasma, show, by using (2.19) or otherwise, that the angle between the wave normal and the normal to the polarisation ellipse for the electric polarisation $\boldsymbol{P}$ is $\arctan\{-\mathrm{i}Y\rho\sin\Theta/(1-X)\}$.

4.4. Two elliptically polarised plane electromagnetic waves of the same frequency, with their wave normals parallel to the z axis, are combined. Their wave polarisations E_y/E_x are respectively $\mathrm{i}a$ and $-\mathrm{i}/a$ where a is real, positive and less than unity. The major axes of their polarisation ellipses are equal. What are the smallest and greatest possible axis ratio (minor axis/major axis) of the polarisation ellipse of the resultant wave and in each case what is the angle between the major axis and the x axis?

4.5. An electromagnetic wave travels through a cold magnetoplasma with its wave normal always parallel to the superimposed magnetic field. The medium attenuates the waves and its thickness is such that the amplitude of the left-handed circularly polarised characteristic wave is reduced by a factor F and that of the right-handed wave is reduced by a factor $\frac{1}{2}F$. What is the state of polarisation of the emergent wave, if the incident wave is (a) linearly polarised (b) unpolarised?

4.6. In an isotropic electron plasma with collisions the refractive index is $\mu - \mathrm{i}\chi$ where μ and χ are real and non-negative. X is held constant and is less than 4/3. Show that if Z is increased from zero, χ at first increases and then attains a maximum value given by $\chi^2 = X^2/\{8(2-X)\}$ when $Z^2 = (4-3X)/(4-X)$. What happens if X exceeds 4/3?

4.7. A homogeneous plasma contains free electrons, and collision damping is negligible. There is a superimposed steady magnetic field which is so strong that it prevents the electrons from moving at right angles to it. The electrons can therefore move only in the direction of the field. A wave of angular frequency ω travels with its wave normal at an angle Θ to the field. It is linearly polarised with its electric vector in the plane containing the wave normal and the direction of the field. Write down the equation of motion of an electron. Hence find the relation between the electric polarisation $\boldsymbol{P}$ and the electric intensity $\boldsymbol{E}$. Show that the refractive index μ is given by:

$$\mu^2 = \frac{1-X}{1-X\cos^2\Theta}$$

where $X = (\omega_N/\omega)^2$ and $\omega_N/2\pi$ is the plasma frequency.

4.8. In a cold electron plasma there is no superimposed magnetic field and electron collisions are negligible. The angular plasma frequency is ω_N and $k = \omega_N/c$. Show that the following inhomogeneous plane waves of angular frequency ω_N are solutions of Maxwell's equations and satisfy the constitutive relations, where A is a constant:

(a) $E_x = \mathrm{i}E_0 \exp\{-A(\mathrm{i}x+z)\}, \quad E_z = E_0 \exp\{-A(\mathrm{i}x+z)\},$
$E_y = \mathscr{H}_x = \mathscr{H}_y = \mathscr{H}_z = 0.$

(b) $E_y = kE_0 \exp\{-A(\mathrm{i}x+z)\}, \quad \mathscr{H}_x = \mathrm{i}AE_0 \exp\{-A(\mathrm{i}x+z)\},$
$\mathscr{H}_z = AE_0 \exp\{-A(\mathrm{i}x+z)\}, \quad E_x = E_z = \mathscr{H}_y = 0.$

In case (b) find the Poynting vector and its time average. (See Budden, 1961a, § 4.9.)

4.9. A homogeneous medium contains N free electrons per unit volume with negligible collision damping, and there is a constant superimposed magnetic field. A linearly polarised plane electromagnetic wave of angular frequency ω travels with its wave normal and its electric vector both perpendicular to the magnetic field. Show that the refractive index for this wave is given by

$$\mu^2 = 1 - X(1-X)/(1-X-Y^2).$$

Show that the electric field has a longitudinal component in quadrature with the transverse component and that the ratio of the longitudinal to the transverse amplitudes is $XY/(1-X-Y^2)$.

Use the Poynting vector to find the instantaneous direction of the energy flow. Hence find the direction of average energy flow when

(a) $X < 1 - Y$ or $X > 1 + Y$, (b) $1 - Y < X < 1 + Y$.
(Maths Tripos 1958, Part III.)

4.10. Discuss the propagation of electromagnetic waves in an isotropic material medium, treating the electric field at each point as the resultant of the electric fields of the incident wave and of the wavelets scattered by elements of volume of the medium.

What extensions of the theory would you expect to have to make to deal with (a) anisotropic media, (b) magnetic rotation of the plane of polarisation?
(Natural Sciences Tripos, 1952, Part II, Physics.)
(See §18.4; Darwin, 1924; Hartree, 1929, 1931b).

4.11. For the two plane progressive waves with the same wave normal direction, and with refractive indices $n = \mu - \mathrm{i}\chi$, in a cold collisionless magnetoplasma, prove that if $Z = Z_t$, (4.27), or equivalently $\Theta = \Theta_t$(4.96), the four curves of μ and χ versus X, (fig. 4.9) have infinite slope where $X = 1$, and the two curves that show how ρ varies in the complex ρ plane are at $\pm 45°$ to the real ρ axis.

4.12. Use the result (e) of problem 3.6 to express Maxwell's equations in principal axis coordinates. Apply them to a plane progressive wave in a homogeneous cold magnetoplasma, with refractive index n and wave normal at an angle Θ to the superimposed magnetic field. Hence derive the dispersion relation in the form

$$\begin{vmatrix} 1+\cos^2\Theta - 2\varepsilon_1/n^2 & -\sin^2\Theta & -\sin\Theta\cos\Theta \\ -\sin^2\Theta & 1+\cos^2\Theta - 2\varepsilon_2/n^2 & -\sin\Theta\cos\Theta \\ -\sin\Theta\cos\Theta & -\sin\Theta\cos\Theta & \sin^2\Theta - \varepsilon_3/n^2 \end{vmatrix} = 0.$$

Show that this is the same as (4.63).

4.13. Show that the solutions of (4.63), giving the two refractive indices for a cold magnetoplasma, may be written

$$n^2 = -\frac{J}{G} - \frac{\frac{1}{2}\varepsilon_3(\varepsilon_1-\varepsilon_2)^2(\varepsilon_1-\varepsilon_3)(\varepsilon_2-\varepsilon_3)}{G\{\frac{1}{2}\varepsilon_3(\varepsilon_1-\varepsilon_2)^2 + GJ\sin^2\Theta\} \pm G^2\{J^2\sin^4\Theta + \varepsilon_3^2(\varepsilon_1-\varepsilon_2)^2\cos^2\Theta\}^{\frac{1}{2}}}.$$

Show that, for an electron plasma, on using (3.50), (3.52) this arrangement of the terms gives the Appleton–Lassen formula (4.47).

4.14. In a collisionless electron plasma there is an extraordinary wave with its wave normal at an angle Θ to the earth's magnetic field. The z axis is parallel to the wave normal and the x axis is coplanar with the wave normal and the earth's field. Show that, when $X = 1$, the fields of this wave satisfy $E_x = \mathscr{H}_y = \mathscr{H}_z = 0$, $\mathscr{H}_x = -E_y$, $E_y/E_z = \mathrm{i}Y\sin\Theta$. Discuss the behaviour of the fields of this wave in the limits (a) $\Theta \to 0$, (b) $Y \to 0$.

4.15. Unpolarised radiation at radio frequencies comes into the earth's ionosphere from the galaxy (at great heights it is reasonable to ignore collisions). Show that within the ionosphere half the energy flux is in the ordinary and half in the extraordinary wave.
(See Budden and Hugill, 1964, especially § 4.)