PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Reflexion Levels and Coupling Regions in a Horizontally Stratified Ionosphere

M. L. V. Pitteway

Phil. Trans. R. Soc. Lond. A 1959 **252**, 53-68
doi: 10.1098/rsta.1959.0014

REFLEXION LEVELS AND COUPLING REGIONS IN A HORIZONTALLY STRATIFIED IONOSPHERE

By M. L. V. PITTEWAY

Mathematical Laboratory, University of Cambridge

(*Communicated by J. A. Ratcliffe, F.R.S.—Received* 27 *February* 1959)

CONTENTS

The propagation of radio waves through a horizontally stratified and slowly varying ionosphere is governed, in the case of oblique incidence, by a quartic equation (Booker 1938). Ray theory breaks down when two roots of this quartic are equal, for then coupling occurs between the characteristic waves, and full wave theory must be used.

This paper is concerned with determining the conditions under which the two roots are equal; it is not concerned with the full wave theory. Values of the plasma frequency, and electron collision frequency, which lead to equal roots, are determined, and are exhibited in a set of curves. A full solution of the 'Booker' quartic is also given for a case of special interest.

It is pointed out that the electric wave-field is unlikely to become very large in a slowly varying ionosphere, so that, if the ionosphere were irregular, scattering cannot be unduly enhanced by a plasma resonance.

1. Introduction

In a recent paper (Budden & Clemmow 1957), the equations governing the propagation of radio waves in a slowly varying and horizontally stratified ionosphere are written in a coupled form, and a ray-theory type of solution, valid where coupling between characteristic waves in the ionosphere is not too large, is given.

It is the purpose of this paper to investigate the levels in the ionosphere at which coupling between the characteristic waves becomes infinite, and 'critical coupling' is said to occur (e.g. Davids & Parkinson 1955). At these levels, ray-theory approximations break down. At normal incidence, the problem of coupling has been fairly thoroughly studied; in this paper, the problem of coupling is extended to include oblique incidence.

For the case of oblique incidence, the Appleton–Hartree equation for the refractive index (Appleton 1932) is replaced by the Booker quartic (Booker 1938) which is used to study propagation in a horizontally stratified slowly varying ionosphere. Coupling is strong near the levels in the ionosphere at which two roots of the Booker quartic become equal, (see, for example, Budden & Clemmow 1957), and so this paper is concerned chiefly with finding the condition satisfied by the ionospheric parameters for the quartic to have two or more equal roots.

The importance of this work lies in the fact that it is very difficult to understand what is happening to a wave propagated through the ionosphere, except in terms of a simple ray-theory picture, and before this picture can be fully understood it is necessary to know both where the reflexion levels lie, and where coupling between the ordinary and extraordinary wave packets is considerable. This is necessary not only when investigating, say, the amplitude of the reflexion of a wave from the ionosphere, or the phase path of a wave packet, but also in a theory of scattering (Pitteway 1958), where it is necessary to find the levels in the ionosphere where fading may be imposed upon a radio wave. In a scattering problem, the scattered waves are oblique even though the original wave may be incident normally, and it is important that the problem of oblique incidence should be fully understood.

It is shown in this paper that a resonance effect associated with very large electric wave-fields, familiar at the level $X = 1$ for an isotropic ionosphere, is present in the Booker quartic (contrary to previous expectations that such a resonance is only shown by a full-wave solution of the differential wave equations), and that when the earth's field is included it is only the extraordinary wave (for frequencies above 1·28 Mc/s), which shows this resonance. Now the extraordinary wave is reflected at an ionospheric level well below the point of resonance, and so an appreciable resonance can only be excited if some extraordinary wave is coupled in by the ordinary wave higher in the ionosphere. At normal incidence, this can happen only if there is some collisional damping present to give the necessary coupling condition, and this in itself will completely suppress any resonance effect (Pitteway 1958).

At oblique incidence, however, it is shown that coupling can occur even though the collisional damping be very small; this is because the wave normals are not constrained to lie parallel to the vertical as they are for normal incidence. It is shown that the critical coupling can occur for as many as seven different values of the ratio between plasma and wave frequencies, even with no collisional damping.

Now ray theory fails when either

(i) two roots of the quartic become equal, or

(ii) one root of the quartic becomes infinite, and in both cases a more detailed study of the differential wave equations is needed. In case (i), the differential equation reduces to Stokes's equation, (except at the penetration of a layer in the ionosphere), and the solutions are Airy integral functions. The point where the two roots are equal is an ordinary point (in the complex plane) of the differential equation, and the equation has no singularity. Thus, although a ray-theory approximation breaks down when two roots are equal, the solutions of the differential equation remain finite, and the electric wave-fields have no singularity.

For case (ii), a number of special cases have been studied, (Försterling & Wüster 1951; Budden 1954), and solution of the differential equations show that in every case the electric wave-field has a singularity. This singularity can be attributed to a plasma oscillation and resonance due to induced space charges (e.g. Herlofson 1951), and we shall therefore speak of case (ii) as a resonance; it is inferred that a full solution of the differential equation will always show a singularity when one of the quartic roots is infinite.

For the case of normal incidence, it has been shown (Budden 1954) that no reflexion takes place at a singularity in the Appleton–Hartree refractive index, and the concept of a 'fourth reflexion condition' for radio waves is discredited. Further, absorption at the singularity is considerable even though the collision frequency may be much less than the wave frequency.

In this paper, it is shown that for propagation at oblique incidence (not in the East–West plane), the Booker quartic has only one infinite root at the singularity, and that there is a separate reflexion level nearby. As normal incidence is approached, this reflexion level moves closer to the singularity, until at normal incidence the quartic has two infinite roots there. In this sense, there is a 'reflexion level' included in the singularity for normal incidence; absorption at the singularity is, however, so large that this reflexion is completely lost (in agreement with Budden 1954). For oblique incidence at low frequencies in the North–South plane, this reflexion level moves well away from the highly absorbing resonance level, and as it is now quite distinct it may be of physical importance.

Section 3 of this paper is devoted to a brief summary of the results of previous work, and the fourth to deriving the condition for the Booker quartic to have two equal roots. Special cases are discussed at length in § 5, and in § 6 graphs of reflexion and coupling levels are given as a function of the angle of incidence. In § 7 the actual solutions of the quartic are discussed for a case of particular interest.

2. Notation

The notation adopted in this paper is essentially the same as that of previous workers. All electric and magnetic quantities are in rationalized units. Some of the more important symbols used are listed below:

x, y, z	Cartesian co-ordinates; z is measured vertically upwards, and incidence is restricted to the x, z plane
θ_i	angle of incidence
S	$\sin\theta_i$
C	$\cos\theta_i$
l, m, n	direction cosines of the earth's magnetic field
ω	angular wave frequency
$\mathbf{B}$	earth's magnetic induction
e, m	charge and mass of the electron
ϵ_0	electric permittivity of free space
N	number of electrons per unit volume
ν	collision frequency of the electrons, giving damping
X	$Ne^2/\epsilon_0 m\omega^2$
Y	$e\mathbf{B}/m\omega$
Z	ν/ω
U	$1-\mathrm{i}Z$
q	a root of the Booker quartic
$\alpha, \beta, \gamma, \delta, \epsilon$	coefficients of the Booker quartic, $F(q) = 0$
A	$\gamma^2-4\alpha\epsilon$
$\mathfrak{A}$	$A+\beta\delta$

$\mathfrak{B}$ $\frac{1}{2}B$, where $B = \beta^2 X^2 (U - X)$

X_c, Z_c values of X and Z which together give coupling between the ordinary and extraordinary waves. (This is not quite the same as the conventional usage, in which $Z_c = |Y_T^2/2Y_L|$; at vertical incidence only, the two definitions of Z_c are equivalent.)

The notation used for labelling figures is outlined in § 6.

3. Previous work

The problem of radio-wave propagation in a given direction through a homogeneous medium has been extensively studied (e.g. Appleton 1932; Ratcliffe 1933).

In a classic paper (Booker 1938), it has been shown that when a plane wave is incident obliquely onto a stratified ionosphere which varies slowly in the vertical direction, its behaviour is represented by a quartic equation, the 'Booker' quartic. The four roots of this quartic give, as a function of height, the behaviour of four distinct waves, corresponding to upgoing and downgoing ordinary and extraordinary waves. These roots can be complex, and when this happens the corresponding wave is attenuated as it travels through the ionosphere.

More recently this work has been extended (Clemmow & Heading 1954; Budden & Clemmow 1957), and it has been shown that the four roots of the quartic are the latent roots of a tensor which governs the wave propagation; the wave equation can be written in a coupled form, and for a slowly varying ionosphere the coupling can be ignored except in regions where two roots of the quartic approach closely. When the coupling can be ignored, the solutions of the tensor wave equation correspond to the four characteristic waves.

If two roots of the quartic are almost equal, ray theory breaks down, and energy is exchanged between the two waves concerned. Thus if the two roots corresponding to the upgoing and downgoing ordinary wave approach, this transfer of energy represents reflexion of the ordinary wave. If the two roots corresponding to the upgoing ordinary and extraordinary wave, say, approach, these two waves cease to travel independently, and an upgoing ordinary wave will excite some extraordinary, and vice-versa.

In general, the coefficients $\alpha, \beta, \gamma, \delta, \epsilon$ of the Booker quartic (see equation (1)) are functions of the electron density variable X, and the collisional damping variable Z, where X and Z are purely real; in the case when $Z = 0$, the coefficients $\alpha, \beta, \gamma, \delta, \epsilon$ are all real. For reflexion, the roots of the quartic are equal only when Z is zero, and for certain critical values of X; if Z never becomes sufficiently small for reflexion to occur, very little energy is returned, and this means that most of the wave is attenuated and lost.

For critical coupling between the ordinary and extraordinary waves, however, it is often necessary to have a non-zero value of Z, written Z_c, before the two roots can approach. The values of q are of necessity complex when a non-zero value of Z is to be introduced, and the condition that both the real and imaginary parts of a pair of roots should be equal gives two equations specifying simultaneously the critical value of X and Z.

For the special case of propagation in the East–West plane, the coefficients of q^3 and q in the quartic, β and δ respectively, both vanish, and the equation becomes a quadratic in the variable q^2. The problem of finding the values of X and Z for which this quartic has equal roots is easy; one set of equal roots occur when ϵ is zero, giving two equal zero roots of the quartic; these correspond to reflexion. Another set are equal when $\gamma^2 = 4\alpha\epsilon$, and this condition corresponds to coupling between the ordinary and extraordinary wave in both the upgoing and in the downgoing directions. There are also two infinite roots of the quartic when $\alpha = 0$; this corresponds to a 'resonance', where the electric wave-fields obtained by solving the differential equations have a singularity.

Normal incidence is included in the above calculation as a special case of East–West propagation. The quartic also reduces to a quadratic in q^2 if the earth's magnetic field is horizontal, and this is another simple case. It is the purpose of this paper to investigate the more difficult problem when β and δ are not zero, and the quartic equation cannot be simplified in this way.

4. General theory

The quartic equation derived by Booker (1938), governing the propagation of a plane radio-wave incident obliquely onto a slowly varying stratified ionosphere, may be written in the notation here adopted

$$F(q) = \alpha q^4 + \beta q^3 + \gamma q^2 + \delta q + \epsilon = 0, \tag{1}$$

where

$$\left.\begin{aligned} \alpha &= U(U^2-Y^2) - X(U^2-n^2Y^2) = U^2(U-X) - Y^2(U-n^2X), \\ \beta &= 2SXlnY^2, \\ \gamma &= -2U[(C^2U-X)(U-X)-C^2Y^2] - XY^2(1-l^2S^2+n^2C^2), \\ \delta &= -2SC^2XlnY^2 = -C^2\beta, \\ \epsilon &= (C^2U-X)[(C^2U-X)(U-X)-C^2Y^2] - S^2C^2Y^2l^2X. \end{aligned}\right\} \tag{2}$$

In this paper, we are concerned with the condition imposed upon the coefficients $\alpha, \beta, \gamma, \delta, \epsilon$ in order that (1) should have two equal roots, $q_1 = q_2$ say. It is well known that (1) will have two equal roots if

$$\left(\frac{\partial F}{\partial q}\right)_{q=q_1} = 0. \tag{3}$$

Differentiating (1), we have

$$4\alpha q^3 + 3\beta q^2 + 2\gamma q + \delta = 0. \tag{4}$$

Thus for the quartic to have two or more equal roots, the values of X and Z, $(U = 1 - iZ)$, must be so chosen that it is possible for q to satisfy (1) and (4) simultaneously. It is convenient to eliminate q between these two equations, obtaining a condition on $\alpha, \beta, \gamma, \delta$ and ϵ alone. This is done by eliminating successively q^4, q^3, q^2 and q to give

$$\begin{vmatrix} 0 & 0 & 4\alpha & 3\beta & 2\gamma & \delta \\ 0 & 0 & \beta & 2\gamma & 3\delta & 4\epsilon \\ 0 & 4\alpha & 3\beta & 2\gamma & \delta & 0 \\ 0 & \beta & 2\gamma & 3\delta & 4\epsilon & 0 \\ 4\alpha & 3\beta & 2\gamma & \delta & 0 & 0 \\ \beta & 2\gamma & 3\delta & 4\epsilon & 0 & 0 \end{vmatrix} = 0, \tag{5}$$

and this determinant of order six can be reduced to the symmetric third-order determinant

$$\begin{vmatrix} 8\alpha\gamma - 3\beta^2 & 6\alpha\delta - \beta\gamma & 16\alpha\epsilon - \beta\delta \\ 6\alpha\delta - \beta\gamma & 4\alpha\epsilon - \gamma^2 + 2\beta\delta & 6\epsilon\beta - \gamma\delta \\ 16\alpha\epsilon - \beta\delta & 6\epsilon\beta - \gamma\delta & 8\epsilon\gamma - 3\delta^2 \end{vmatrix} = 0. \tag{6}$$

The central term in (6) contains the term $(\gamma^2-4\alpha\epsilon)$, which is a factor of some importance in the East–West case when β and δ are both zero, and it is convenient to write

$$\begin{aligned} A &= \gamma^2-4\alpha\epsilon \\ &= X^2Y^2[Y^2(1-l^2S^2-n^2C^2)^2+4(U-X)\,(Ul^2S^2+Un^2C^2-n^2X)]. \end{aligned} \tag{7}$$

Another simple combination of the quartic coefficients occurring in the expansion of (6) is defined by

$$\begin{aligned} B &= \alpha\delta^2+\epsilon\beta^2-\beta\gamma\delta \\ &= \beta^2X^2(U-X). \end{aligned} \tag{8}$$

The terms in the expansion of (6) may be grouped to give, at length,

$$16(A+\beta\delta)^3-16\gamma^2(A+\beta\delta)^2+144\gamma B(A+\beta\delta)+108B^2-128\gamma^3B = 0.$$

Write $\mathfrak{A}$ for $(A+\beta\delta)$, $\mathfrak{B}$ for $\frac{1}{2}B$, (B contains a factor 4 through the term β^2), and divide by 16. We have

$$\mathfrak{A}^3-\gamma^2\mathfrak{A}^2+18\gamma\mathfrak{B}\mathfrak{A}+27\mathfrak{B}^2-16\gamma^3\mathfrak{B} = 0, \tag{9}$$

and this is the required expression, which gives the condition that the Booker quartic should have two equal roots. It is especially convenient in that only three separate functions of X and Z, $\mathfrak{A}$, $\mathfrak{B}$ and γ are involved, instead of the five expressions $\alpha, \beta, \gamma, \delta$ and ϵ originally present.

The left-hand side of equation (9) contains the obvious factor X^4Y^4, as each term contains this factor. This leads at once to the trivial cases:

(i) $X = 0$. In free space below the ionosphere there is no distinction between ordinary and extraordinary waves in the upward or in the downward directions, and these waves will have equal values of q.

(ii) $Y = 0$. Again, there is no distinction between ordinary and extraordinary waves in the isotropic ionosphere resulting when $Y = 0$, and the corresponding roots of the quartic will be equal for all values of X and Z.

We can now remove this factor X^4Y^4, which shows an improvement on the determinant forms (5) and (6) in which the factor is certainly not obvious. Other conditions for equal roots are obtained by searching numerically for the zeros of the remaining function, and this has been done on a digital computer. But first it is of interest to consider some special cases.

5. Special cases

In the previous section, a function of X and Z, given by equation (9), has been derived, the zeros of which give the conditions for the Booker quartic (1) to have equal roots. For the case when Z is zero (and U is thus unity), the function is purely real—a polynomial function of X—the zeros of which give the reflexion levels in the ionosphere. Some of these zeros will correspond to real values of X, and others with complex values of X will not be considered here. When Z is non-zero, the function is complex, and a numerical search is necessary to find values of X and Z which together make both the real and imaginary parts of (9) zero simultaneously.

In this section, partly as a check on the analysis, it is of interest to consider certain special cases.

(*a*) *East–West*

For East–West propagation, the direction cosine l of the earth's magnetic field vanishes, and so β and δ both vanish, leaving a quadratic equation in q^2. In (9), $\mathfrak{A} = A$, and $\mathfrak{B}$ vanishes. The condition for equal roots becomes

$$A^2(A-\gamma^2) = 0,$$

i.e.

$$-4\alpha\epsilon(\gamma^2-4\alpha\epsilon)^2 = 0,$$

thus either (i) $\epsilon = 0$, or (ii) $\alpha = 0$, or (iii) $(\gamma^2-4\alpha\epsilon^2)^2 = 0$.

Case (i) $\epsilon = 0$

Since $\delta = 0$, this gives equal roots at $q = 0$ for $Z = 0$, $X = C^2$ (ordinary reflexion level), or $Z = 0$ and

$$(C^2-X)(1-X) = C^2Y^2,$$

i.e.

$$X = \tfrac{1}{2}\{1+C^2\pm\surd(S+4C^2Y^2)\},$$

(the two extraordinary reflexion levels). In the case when $Y > 1$, one of the extraordinary reflexion levels occurs for negative (non-physical) values of X, as is well known.

Case (ii). $\alpha = 0$

Since $\beta = 0$, two roots of the quartic are infinite when $\alpha = 0$, and as the differential equation in the electric fields shows a singularity at this point, the electric wave-field becomes infinite, and a resonance occurs. (In the more general case when β is not zero, the zero of (9) moves away from the point where $\alpha = 0$; one of the roots only goes off to infinity, and there is an extraordinary reflexion level nearby.) This resonance with $\alpha = 0$ requires $Z = 0$, $X = (1-Y^2)/(1-n^2Y^2)$.

Case (iii). $(\gamma^2-4\alpha\epsilon)^2 = 0$

This is more interesting, and gives for the East–West case the condition for critical coupling between the ordinary and extraordinary waves, an extension to oblique incidence of the work of Davids & Parkinson (1955). We require A to be zero for the case when $l = 0$ and the factor X^2Y^2 is discarded, i.e. from (7)

$$Y^2(1-n^2C^2)^2+4(U-X)\,n^2(C^2U-X) = 0.$$

Equating real and imaginary parts,

$$Y^2(1-n^2C^2)^2+4(1-X)\,n^2(C^2-X)-4n^2C^2Z^2 = 0,$$

and

$$Z[4n^2(C^2-X)+4n^2C^2(1-X)] = 0.$$

We have either
$$X_c = \frac{2C^2}{1+C^2};\quad Z_c = \sqrt{\left(\frac{Y^2(1-n^2C^2)^2}{4n^2C^2}-\frac{S^4}{(1+C^2)^2}\right)},$$

or
$$X_c = \frac{1}{2}\left[1+C^2\pm\sqrt{\left(S^4-\frac{Y^2(1-n^2C^2)^2}{n^2}\right)}\right];\quad Z_c = 0.$$

It is interesting to observe that for oblique incidence it is possible to have coupling between the ordinary and extraordinary wave-fields, even though Z may be zero and no collisional damping is present. This does not occur for vertical incidence, where the values of q are the same as μ given by the Appleton–Hartree formula (Appleton 1932). This is because, at normal incidence, the ordinary and extraordinary wave normals are everywhere parallel, whereas, at oblique incidence, in general they are not. Some curves are plotted for East–West coupling at a later stage in this paper, where the manner in which the curves of X_c for zero and non-zero values of Z_c link is illustrated.

It is interesting to see that for the East–West case, the critical coupling term $A = 0$ occurs as a factor A^2. This is because, for this special case, critical coupling between the upgoing ordinary and extraordinary waves occurs at the same values of X_c and Z_c as the downgoing waves, so the factor A occurs twice.

(b) *Normal incidence*

This may be regarded as a special case of East–West propagation, and the results are well known. We replace S by 0 and C by 1 in the East–West work, and we have reflexion levels for $X = 1$ (ordinary) and for $X = 1 \pm Y$ (extraordinary), resonance at

$$X = (1-Y^2)/(1-n^2Y^2),$$

and critical coupling between the ordinary and extraordinary waves when $X_c = 1$ and $Z_c = Y(1-n^2)/2n$. There is no critical coupling with $Z_c = 0$ for real values of X_c.

In interpreting the graphs drawn for the general case at a later stage in this paper, it is helpful to understand what happens to a wave-packet incident normally, so that the different regions of propagation and reflexion may be identified. This work is familiar to most radio physicists, and is given in some detail in Ratcliffe (1959).

(c) *Earth's field horizontal*

On the geomagnetic equator, the n-direction cosine of the earth's field vanishes, and β and δ are again zero. Reflexion levels occur where $\epsilon = 0$, and critical coupling occurs where

$$Y^2(1-l^2S^2)^2 + 4Ul^2S^2(U-X) = 0,$$

i.e.

$$X_c = 2, \quad Z_c = \sqrt{\left(\frac{Y^2(1-l^2S^2)^2}{4l^2S^2} - 1\right)},$$

or

$$X_c = \frac{Y^2(1-l^2S^2)}{4l^2S^2} + 1, \quad Z_c = 0.$$

From this, and from case (a), we can see that there can be no ordinary–extraordinary coupling for East–West propagation at the geomagnetic equator where $l = n = 0$, and $A = X^2Y^4$.

Resonance occurs where $Z = 0$ and $X = (1-Y^2)$.

(d) $Y = 0$, *no field*

In the limiting process $Y \to 0$ it is useful to discard the trivial root Y^4 in (9), leaving

$$-64U^2(U-X)^4(C^2U-X)^2X^4[U(l^2S^2-n^2C^2)+n^2X]^2 = 0;$$

the factor $(U-X)^4$ gives the Herlofson resonance level at $X = 1$, $Z = 0$ and $(C^2U-X)^2$ gives the reflexion level at $X = C^2$, $Z = 0$. The last factor gives the level at which critical coupling might be expected for small values of Y.

Discarding the term in Y^4 before allowing Y to become small has led to a very important point. With $Y = 0$, the Booker quartic reduces to

$$(U-X)(Uq^2-C^2U+X)^2 = 0,$$

giving
$$q^2 = C^2 - \frac{X}{U}. \tag{10}$$

This shows no singularity at $X = 1$, $Z = 0$, and it is natural to think that the singularity in E_x and E_z at $X = 1$, $Z = 0$ must arise from some feature of the full-wave solution. But the limiting process $Y \to 0$ shows that the Booker quartic really contains this singularity. This point may be brought out by writing the Booker quartic, from (1) and (2), in the form

$$(U-X)[U(q^2-C^2)+X]^2 = Y^2(q^2-C^2)[U(q^2-C^2)+X\{1-(lS+nq)^2\}]. \tag{11}$$

In the special case when $X = 1$ and $Z = 0$, the left-hand side of (11) vanishes. We may divide the right-hand side by Y^2, and the roots are

$$q = \pm C \quad \text{or} \quad \frac{S}{(1-n^2)}(ln \pm im).$$

For any value of Y, however small, the solutions of the quartic will be $\pm C$, and two other roots, which depend upon the direction cosines of Y, and are indeterminate when $Y = 0$.

In the special case when $Y = 0$, the right-hand side of (11) vanishes, the factor $(U-X)$ divides out, and we have equation (10). Now if we set $X = 1$ and $Z = 0$, we obtain $q^2 = -S^2$, two equal roots at $+iS$ and two equal roots at $-iS$. Clearly, the limiting processes $Y \to 0$, and $(U-X) \to 0$, may not be reversed, and the resonance at $X = 1$ really is concealed in the Booker quartic.

(*e*) *North–South*

For the case of propagation in the North–South plane, the direction cosine m vanishes, so that $l^2+n^2 = 1$. It can easily be shown that (9) vanishes in the North–South case at all angles of incidence at the point where $X = 1$ and $Z = 0$, a well-known special case. It is of some interest to remove the factor $(U-X)$ from (9), and to investigate the condition that there should be a second solution of (9) at $X = 1$, $Z = 0$; this can also be done quite simply, and it is found that a second zero of (9) occurs if

(i) $C^2 = 0$, grazing incidence, or
(ii) $C^2 = n^2$ (labelled B on graphs), or
(iii) $S^4(1-Y^2)+2l^2S^2Y^2-l^4Y^2 = 0$,

$$S = \pm l\sqrt{\frac{Y}{Y\pm 1}} \quad \text{(labelled } A_1, A_2\text{)}.$$

This is in agreement with previous workers, and provides a useful check on the numerical work for the North–South case. Other solutions of (9) for North–South propagation are complicated, and are best found numerically.

(*f*) *Grazing incidence*

For grazing incidence, it can be shown that the function (9) has no simple roots, except one at $X = 1$, $Z = 0$ corresponding to the extraordinary reflexion level. This justifies the numerical methods of the next section, for it follows that in general (9) will have no useful factors, and no useful result can be expected from a detailed algebraic manipulation of (9).

6. Numerical work

Before applying numerical methods to search for the zeros of (9), it is useful to consider the symmetry of the problem, with a view to avoiding unnecessary repetition of calculations. For example, the coefficients α, β, γ, δ, ϵ involve Y only through the square Y^2, and are invariant to a change of sign of Y. This means that we may change the sign of Y without altering the solutions of the quartic.

If θ_i, the angle of incidence, is changed in sign, S changes sign; the coefficients β and δ change sign, while α, γ and ϵ are unaltered. Thus the signs of the four roots of the quartic are reversed, and two equal roots remain equal. It follows that the values of X and Z which make (9) vanish are unaltered, and this is obvious in (9), for S is involved only as S^2.

Further symmetry is introduced by the occurrence of l as l^2 only in (9), for l also affects the sign of just β and δ; thus a change of the plane of incidence which just changes the sign of l will not alter the zeros of (9), and the reflexion and coupling regions will be the same, for example, in all four possible directions midway between North–South and East–West, that is,

$$\text{NE} \to \text{SW}, \quad \text{SW} \to \text{NE}, \quad \text{NW} \to \text{SE} \quad \text{and} \quad \text{SE} \to \text{NW}.$$

In preparing numerical work, full advantage was taken of this symmetry, and the program of work included an investigation in the E → W plane (hence also W → E), N → S (hence, also S → N), and the midway general azimuth case.

A program for EDSAC II, the digital computer in the Mathematical Laboratory, University of Cambridge, has been devised to search for the zeros of the function given by (9). This program set the computer to read in an approximate solution, X and Z, and the machine then proceeded to calculate the function for this X and Z, and also at two neighbouring values of X and Z. When this had been done, the computer used the values at the three points to calculate approximate values for the gradients, and thus an improved value for X and Z, this whole cycle taking about one-tenth of a second. Some ten or twenty cycles were usually sufficient to define X and Z with five-figure accuracy, (though occasionally, when two zeros occurred for nearly equal values of X and Z, longer was necessary). The machine then punched out the value of X and Z derived, solved the Booker quartic as a check, (and to identify whether the zero was a reflexion level or a coupling region), and then proceeded to the next calculation.

Some of the results of this work are shown graphically in figures 1 to 10, pp. 67 and 68, where, for different values of Y, and for different planes of incidence, values of X and Z which satisfy the condition (9) are plotted as functions of θ_i, the angle of incidence. Values of X are shown by a continuous line, and values of Z by a broken line. The curves are labelled according to the following code:

σ Ordinary reflexion level; this gives the value of X at which the wave packet with ordinary polarization will be reflected, and it is understood that the value of Z associated with this X is zero. In all cases, σ drops from $X = 1$ at normal incidence to $X = 0$ at grazing incidence.

E Extraordinary reflexion level; this also must be associated with $Z = 0$. For $Y < 1$, there are two curves for E reflexion, as would be expected from the normal incidence case. For $Y > 1$, there is only one curve.

R Again associated with $Z = 0$, this is an *E*-reflexion level occurring near the point where $\alpha = 0$ and one root of the quartic is infinite. *R* stands for resonance, because for normal incidence, or East–West propagation, this occurs at exactly the point where $\alpha = 0$ and resonance occurs. It is only for oblique incidence out of the East–West plane and away from the geomagnetic equator that the *R* curve is distinguished from the resonance at $\alpha = 0$.

X_c, Z_c Values of *X* and *Z* at which 'critical coupling' occurs, that is, coupling between ordinary and extraordinary waves. For the East–West special case the upgoing waves couple together at the same values of *X* and *Z* as the downgoing waves, and there is only one X_c, Z_c curve. More generally, there are two sets of curves, X_{c1}, Z_{c1} and X_{c2}, Z_{c2}, one giving critical coupling between the upgoing ordinary and extraordinary waves, and one between the downgoing waves. It has already been mentioned that when the sign of θ_i is changed, the sign of β and δ, and hence of all the roots of the quartic, are changed; thus suppose X_{c1} and Z_{c1} give critical coupling for the upgoing wave; if the sign of θ_i is changed, it will be X_{c2} and Z_{c2} that give critical coupling for the upgoing wave. This is clearly illustrated in, say, figure 2, where, remembering that the curves repeat symmetrically for negative values of θ_i, the crossing of X_c and Z_c at the origin is clearly evident.

X_c' A value of *X* giving critical coupling for the case when $Z_c = 0$. In some cases, as θ_i increases from zero, Z_c decreases and vanishes at a critical angle; at this point, the curve for X_c splits into two separate X_c' curves. In figure 1, a small insert shows, on a greatly enlarged scale, how this split occurs.

Graphs are drawn for the cases when $Y = 0{\cdot}5$, 2 and 1 for E–W, N–S, and general azimuth midway between, and also in the E–W case for $Y = 0{\cdot}1$ to illustrate the transition to the isotropic case when $Y = 0$. Figure numbers are allocated according to the scheme:

Figure	Azimuth	*Y*	
1	E–W	$Y = 0{\cdot}5$	'angle of dip' $= 60°$, so that $n = -\frac{1}{2}\sqrt{3}$, in all cases.
2	N–S		
3	45°		
4	E–W	$Y = 2{\cdot}0$	
5	N–S		
6	45°		
7	E–W	$Y = 1{\cdot}0$	
8	N–S		
9	45°		
10	E–W	$Y = 0{\cdot}1$	

In the North–South figures it is interesting to see how the well-known coupling level, $X = 1$, $Z = 0$, can represent σ, X_c' and *R* labels. For example, consider figure 5; the root at $X = 1$ represents ordinary wave reflexion until θ_i reaches the value at A_1 given by $S = \pm l\sqrt{\{Y/(Y+1)\}}$, where it becomes a critical coupling level with $Z_c = 0$ (X_c'). The second X_c' curve crosses the line $X = 1$ at the point where $C^2 = n^2$, labelled *B*, and then at A_2, where $S = \pm l\sqrt{\{Y/(Y-1)\}}$, another transition occurs, and $X = 1$ represents an *R* reflexion. For a value of *Y* less than unity, (see figure 2), this latter transition does not occur.

In most of the figures it can be seen that sometimes a value of X_c goes down to zero (and would go negative had this been permitted), while the value of Z_c necessary for critical coupling has a singularity. It can easily be shown from (9) that this must occur where $l^2S^2 = n^2C^2$, and propagation is in a direction perpendicular to Y. For the East–West figures, this occurs where $\theta_i = 90°$, for North–South where $\theta_i = 60°$, (since $n = -\frac{1}{2}\sqrt{3}$), and midway where θ_i is equal to 67° 48′ approximately. These points are marked D on the curves. The effect of critical coupling at large values of Z_c for small values of X_c will be very important in determining the limiting polarization of a downgoing radio wave, as low in the ionosphere the ordinary and extraordinary wavelength is very similar, and even weak coupling can produce a cumulative effect (considered, for vertical incidence only, by Budden 1952).

With reference to the possibility of a resonance in the scattering of radio waves (Pitteway 1958), it is interesting to view the possibility of very large electric wave-fields in the light of this work. Take, for example, a frequency near 2 Mc/s. Except at very oblique incidence, the extraordinary wave is reflected low in the ionosphere well below any resonance level, yet it is the extraordinary wave which might be expected to give large scattering wave-fields higher up. The extraordinary wave can therefore only be present in any appreciable quantity if it is coupled in by some value of collisional damping approaching Z_c, and this will normally require a value of Z of order 0·1. Now it has already been shown for the isotropic ionosphere that so large a value for Z will heavily suppress any resonance, and it seems extremely likely that that will apply here also. Thus the only possibility of obtaining resonance scattering will occur at angles of beyond 30° for East–West, (see figure 1), or at any rate beyond about 15° (see figure 2). This is very important, for it means that scattering at oblique angles may be encouraged, (it must be remembered that the angle of the scattered wave plays a reciprocal part with the angle of the incident wave), and a tendency to form a conical angular power spectrum with a corresponding oscillatory correlogram at the ground may result. On the other hand, it must be remembered that absorption at the resonance level will be very high even with Z small (Budden 1954), and this will also reduce the scattering effect.

In concluding this section, it is interesting to examine figure 10, which illustrates the transition to the isotropic case when Y is very small. In the isotropic case there is one reflexion level at $X = C^2$, $Z = 0$, and one resonance level at $X = 1$, $Z = 0$, (which is easily missed in the limiting process $Y \to 0$). In figure 10, it can be seen (though the separate branches are not easily distinguished in the graphical representation), that the reflexion level has been resolved by the presence of the small Y into one E-reflexion level, the σ-reflexion level, and an X_c' coupling; the resonance level at $X = 1$ has been split into the second E-reflexion level, the second X_c' curve, and the R curve (here constant at the level where $\alpha = 0$). This illustrates clearly that the resonance at $X = 1$ really is present in the Booker quartic, and is not just a property of the full-wave solution of the differential equations, though great care is necessary in the limiting processes $Y \to 0$ and $X/U \to 1$, which cannot be reversed.

7. Numerical example

In conclusion, in order to illustrate some of the features of the previous work, it is interesting to examine the roots of the quartic themselves as a function of X for a special case. The case chosen corresponds to figure 2 with $\theta_i = 45°$. Propagation is from North to South, with a frequency of 2·56 Mc/s, so that $Y = 0·5$. Curves of q against X are plotted in figure 11 (p. 68) for zero collisional damping, so that $Z = 0$ everywhere.

The coupled wave equations (Budden & Clemmow 1957) break down at no less than seven distinct values of X, six of which give equal roots of the quartic; the seventh corresponds to $\alpha = 0$, where one of the four roots goes off to infinity. As X increases from zero, the successive breakdowns occur at levels given by $\theta_i = 45°$ on figure 2, and these are:

(i)	$X = 0{\cdot}297$ labelled	E;
(ii)	$X = 0{\cdot}577$	σ;
(iii)	$X = 0{\cdot}876$	R;
(iv)	$X = 0{\cdot}923$	$\alpha = 0$, one root of quartic infinite;
(v)	$X = 0{\cdot}9987$	X'_c;
(vi)	$X = 1$	X'_c;
(vii)	$X = 1{\cdot}161$	E;

and these points are marked in figure 11, where q is plotted against X. The real parts of the four values of q are denoted by a continuous line, and the imaginary parts by a broken line. The curves are labelled O or E for ordinary or extraordinary, and U or D for upgoing or downgoing, where appropriate.

For small values of X, the four values of q are all real, corresponding to propagating ordinary and extraordinary wave packets. At the level $X = 0{\cdot}297$ the first E reflexion occurs, and the corresponding roots of the quartic join, and then split into a conjugate complex pair, the real part of which is marked simply E. Provided X is supposed to increase slowly and steadily, the full solution of the wave equations would behave like an Airy function in this region in the usual manner at a reflexion level.

A little higher in the ionosphere at the point where $X = 0{\cdot}577$ the ordinary wave packet is also reflected, and the splitting of two roots into a conjugate complex pair is repeated. Now both the ordinary and extraordinary waves become evanescent, and unless the gradient of X is steep there would in fact be little penetration further into the ionosphere.

The next point of interest is at $X = 0{\cdot}876$, where the R level occurs. Here the two roots corresponding to the extraordinary wave, which are conjugate complex, join together and split into two purely real values, a kind of inverted reflexion level. The behaviour of a wave function would again be like an Airy function, only reversed. A little higher again, where $X = 0{\cdot}923$, one of the two real roots passes through infinity, and returns on the positive side.

In the East–West case, the level R, and this singularity where $\alpha = 0$, would be superimposed. Both E roots would become infinite together, and would come back real, the process of reflexion having been absorbed by the singularity. Here, the reflexion level is moved away from the singularity and is clearly resolved. In figure 12*a* (p. 68), this part of the graph is repeated on an exaggerated X scale for the extraordinary wave only, in order that this reflexion behaviour may be made clear.

In this particular case, there are values of X corresponding to X'_c, which occur next as a wave packet penetrates the ionosphere. But first it is better to change to a level high in the ionosphere where $X = 1{\cdot}5$ say, so that the other E-reflexion level may be examined. At this high level, both the ordinary and the extraordinary roots of the Booker quartic occur in conjugate complex pairs, and the corresponding wave packets are evanescent in behaviour. This is because the extraordinary wave packet, (which becomes purely propagative at the level $X = 0{\cdot}876$), has again been reflected, this time at the second E level where $X = 1{\cdot}161$.

This is perfectly consistent with the behaviour at normal incidence familiar from the Appleton–Hartree refractive index.

Now we consider the two X_c' levels, here occurring close together at the levels where $X = 0{\cdot}9987$ and $X = 1$, respectively. Since these two levels are so close together, it is helpful to examine the q curves on a greatly exaggerated X scale in figure 12*b*. The behaviour of the downgoing extraordinary wave is straightforward, (see the curve DE on figure 11), so this has been omitted from 12*b*.

The upgoing extraordinary wave couples with the two ordinary waves in a rather complicated manner, and in fact consistent labelling of the two curves requires two upgoing extraordinary waves, and only one ordinary wave, for the range $0{\cdot}9987 < X < 1$. There is, of course, no singularity in the differential wave equation at either point, and the electric wave-fields are again finite with an Airy function behaviour.

Had these curves been repeated for non-zero values of Z, it would have been found that the downgoing extraordinary wave would be coupled in where $X = 0{\cdot}341$ and $Z = 0{\cdot}875$, the requisite value of X_c and Z_c for this case. This is, therefore, an eighth point in the XZ plane where the coupled wave equations would break down and the coupling become large. Many such cases have been investigated on the EDSAC, which can find a set of solutions for a quartic with complex coefficients, to five-figure accuracy, in about two seconds even with roots close together, but these are outside the scope of this paper, most of the new features being already present in the special case of figure 11.

I would like to express my thanks to Dr K. G. Budden and Mr J. A. Ratcliffe, F.R.S., for many helpful discussions, suggestions and criticisms. I am indebted to the Director and Staff of the Mathematical Laboratory, University of Cambridge, for permission to use the EDSAC, and to Mrs M. L. V. Pitteway for assistance with heavy algebra. I was in receipt of a maintenance grant from the Department of Scientific and Industrial Research for the period during which the work was carried out.

References

Appleton, E. V. 1932 *J. Instn Elec. Engrs*, **71**, 642.
Booker, H. G. 1938 *Phil. Trans.* A, **237**, 411.
Budden, K. G. 1952 *Proc. Roy. Soc.* A, **215**, 215.
Budden, K. G. 1954 *The physics of the ionosphere*, p. 320. London Physical Society.
Budden, K. G. & Clemmow, P. C. 1957 *Proc. Camb. Phil. Soc.* **53**, 669.
Clemmow, P. C. & Heading, J. 1954 *Proc. Camb. Phil. Soc.* **50**, 319.
Davids, N. & Parkinson, R. W. 1955 *J. Atmos. Terr. Phys.* **7**, 173.
Försterling, K. & Wüster, H. O. 1951 *J. Atmos. Terr. Phys.* **2**, 22.
Herlofson, N. 1951 *Ark. Fys.* **3**, 247.
Pitteway, M. L. V. 1958 *Proc. Roy. Soc.* A, **246**, 556.
Ratcliffe, J. A. 1933 *Wireless Engng*, **10**, 354.
Ratcliffe, J. A. 1959 *Magneto-ionic theory*. Cambridge University Press.

FIGURE 1. Reflexion levels and coupling regions as a function of the angle of incidence for East–West propagation. Y has been given the value 0·5, corresponding to a radio wave of frequency 2·56 Mc/s. The curves are labelled according to the code of § 6, and an insert shows how X_c can split into two values X_c'. The 'angle of dip' is 60° in this and in the subsequent figures.

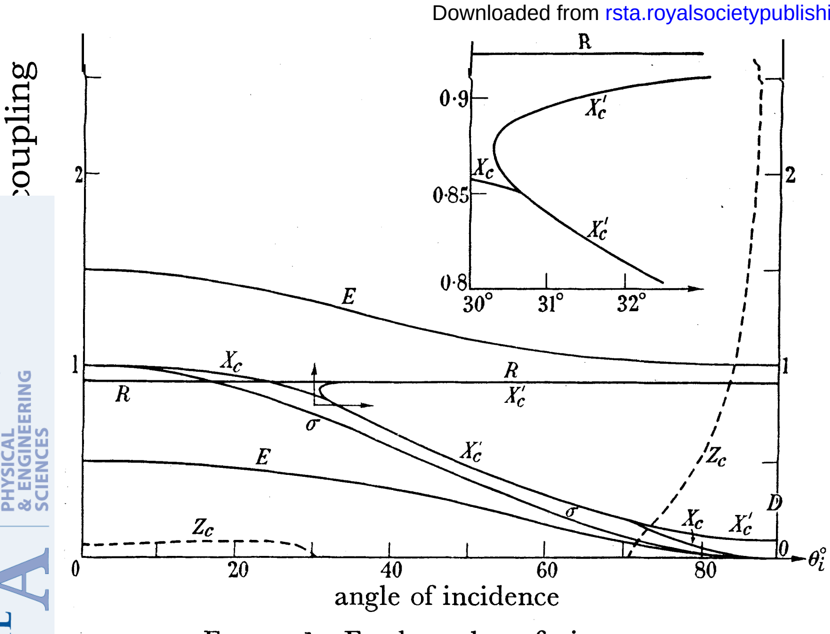

FIGURE 1. For legend see facing page.

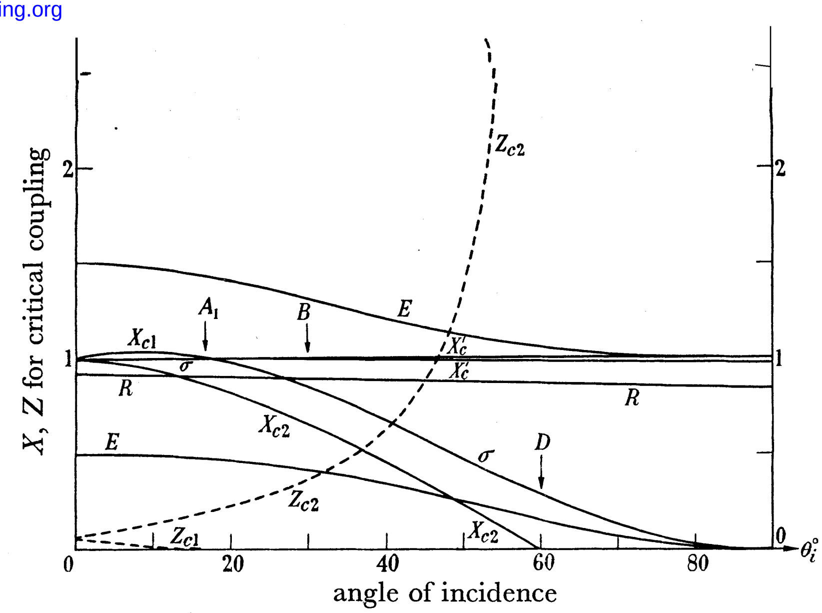

FIGURE 2. $Y = 0{\cdot}5$ (2·56 Mc/s), N–S propagation.

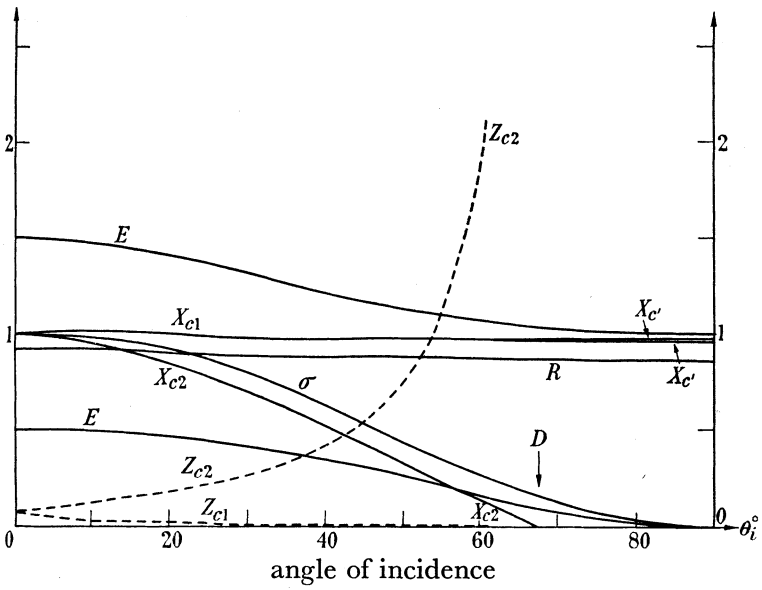

FIGURE 3. $Y = 0{\cdot}5$ (2·56 Mc/s), general azimuth.

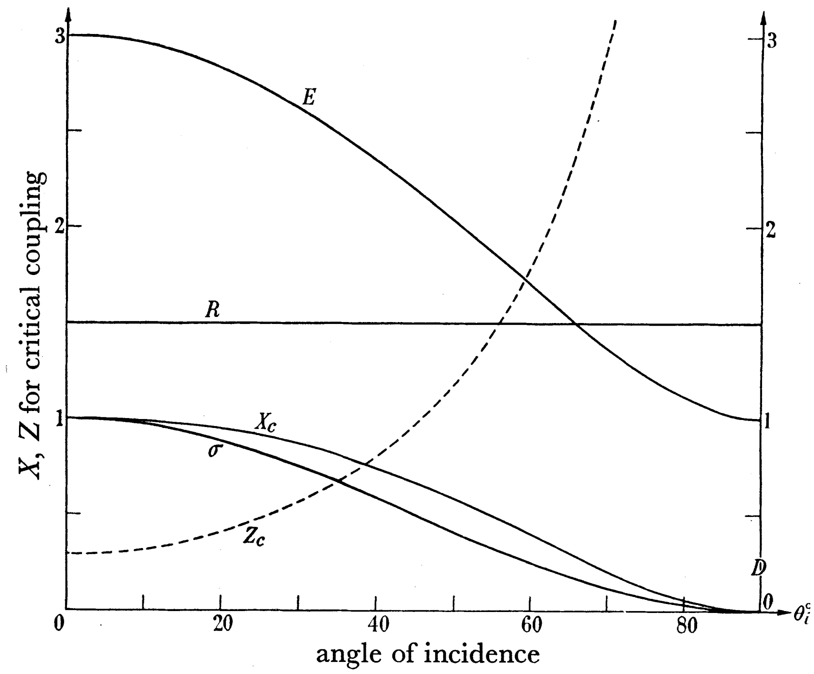

FIGURE 4. $Y = 2{\cdot}0$ (0·64 Mc/s), E–W propagation.

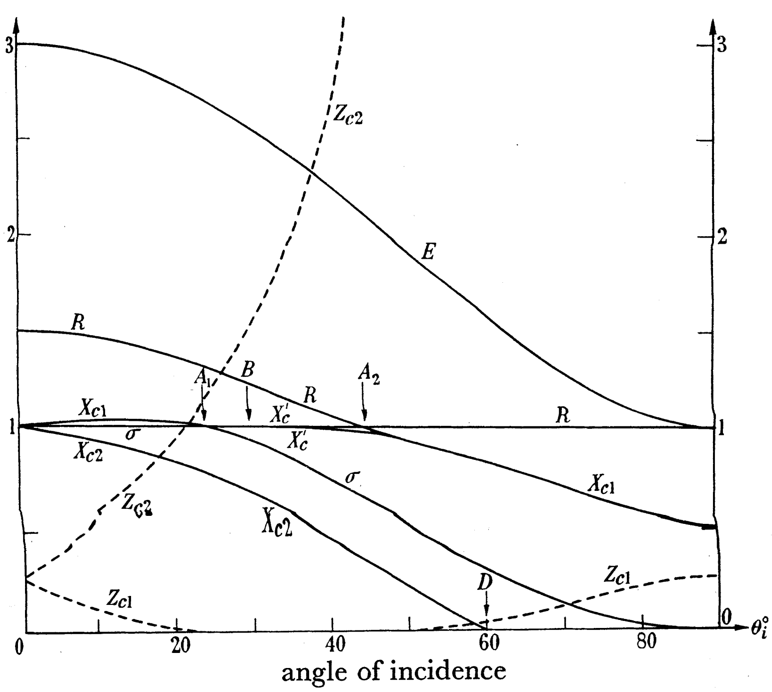

FIGURE 5. $Y = 2{\cdot}0$ (0·64 Mc/s), N–S propagation.

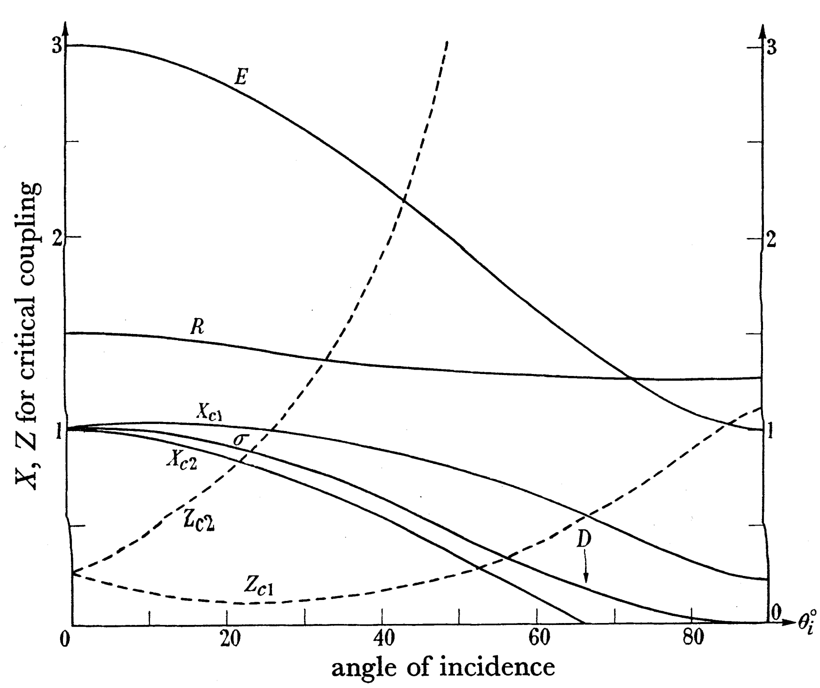

FIGURE 6. $Y = 2{\cdot}0$ (0·64 Mc/s), general azimuth.

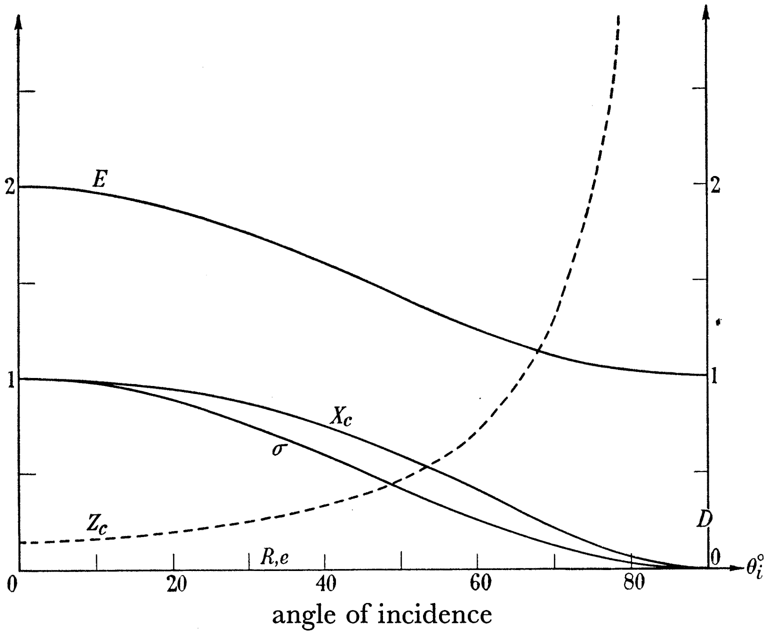

FIGURE 7. $Y = 1{\cdot}0$ (1·28 Mc/s), E–W propagation.

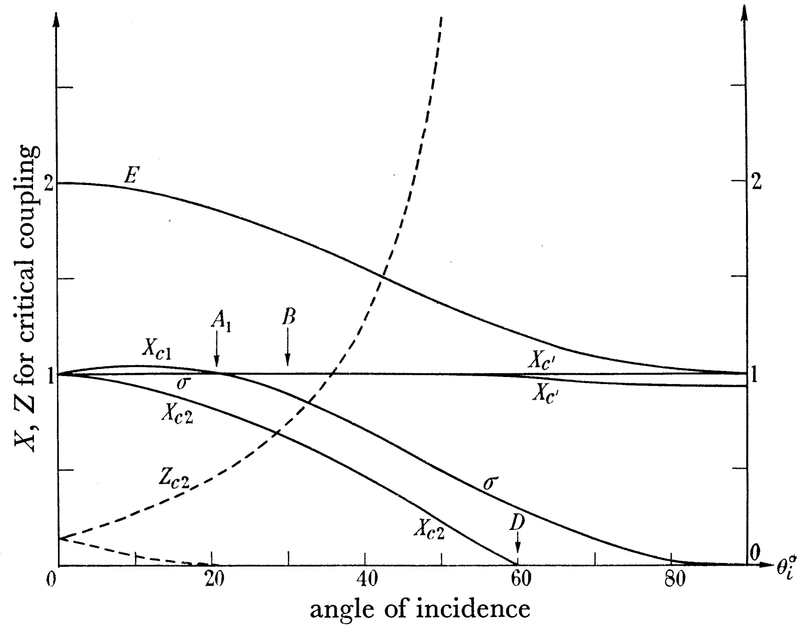

FIGURE 8. $Y = 1{\cdot}0$ (1·28 Mc/s), N–S propagation.

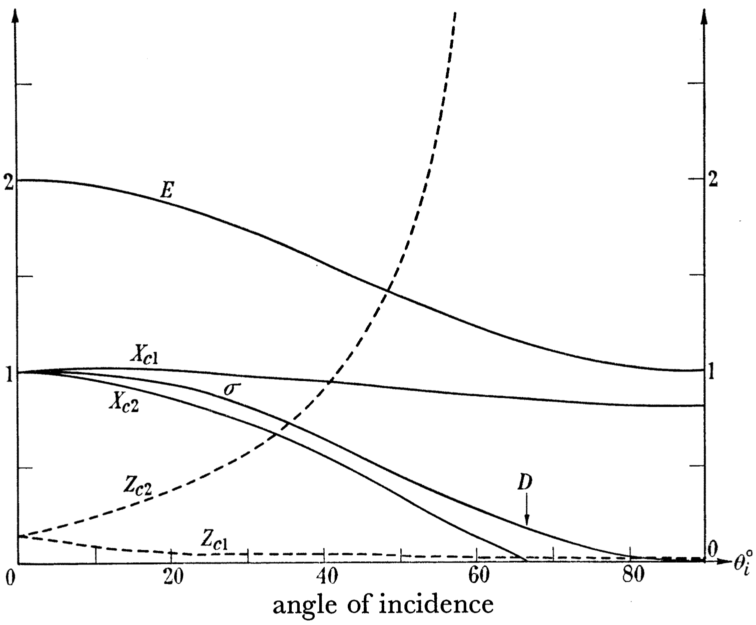

FIGURE 9. $Y = 1{\cdot}0$ (1·28 Mc/s), general azimuth.

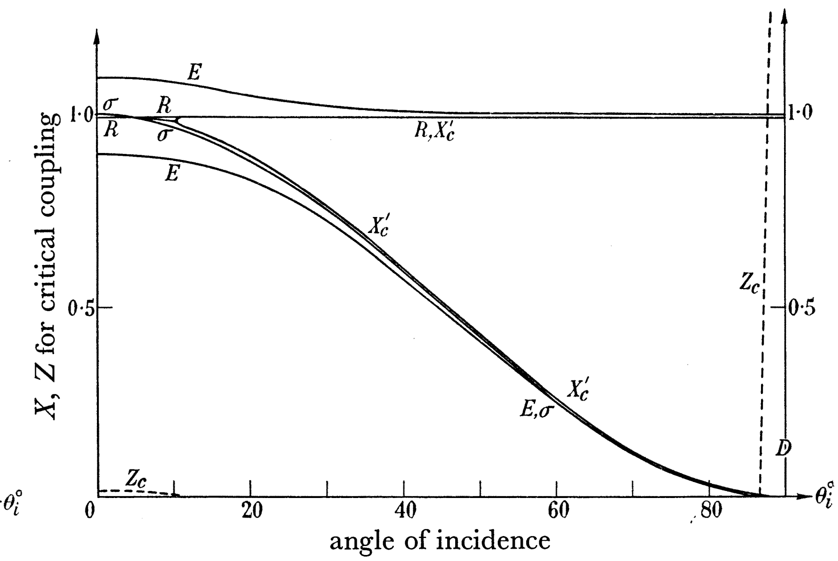

FIGURE 10. $Y = 0{\cdot}1$ (12·8 Mc/s), E–W propagation, illustrating the transition to the isotropic case when $Y = 0$.

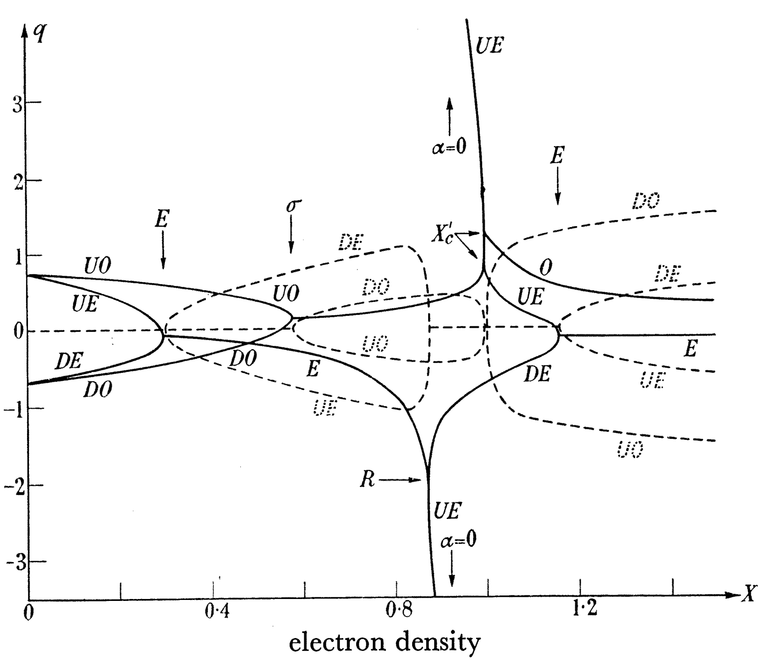

FIGURE 11. The four roots of the Booker quartic as a function of X; $Z = 0$. Propagation is from North to South at an angle of incidence of 45°, Y being 0·5 as in figure 2.

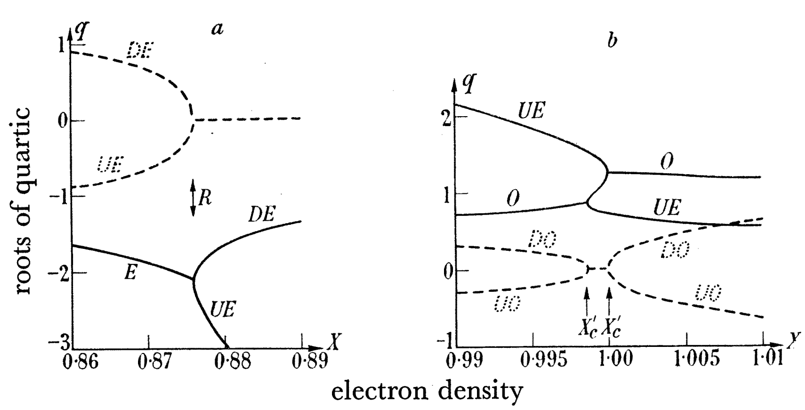

FIGURE 12. *a*, Behaviour of the roots of the Booker quartic in the region of R. *b*, Behaviour of the roots of the Booker quartic in the region of X_c'.