Preface

Electromagnetic theory is fundamental to the whole of electrical and electronic engineering. As such it should surely be an essential part of the professional knowledge of all who call themselves electronic engineers. Yet it is a common complaint among teachers of the subject that students cannot be persuaded to take it seriously perhaps because of their obsession with digital electronics. It seems to me that this is very regrettable. Advances in high-speed digital electronics and in opto-electronics will present problems which cannot be solved without an understanding of electromagnetic theory. The EEC directive on electromagnetic compatibility to be adopted in 1992 will likewise demand a knowledge of fundamental principles. For these reasons it is vital that all students of electrical and electronic engineering should gain a basic knowledge of electromagnetics.

My own experience of grappling with problems in the engineering applications of electromagnetic theory has convinced me that the subject is usually presented in an over-mathematical way. This may be an additional reason why students find it unattractive. In this book I have introduced the subject in a physical and intuitive way making use of elementary mathematics for the most part. The emphasis is on the physical understanding which is the basis for solving problems. Those who eventually need to understand the full mathematical treatment will find that this book provides a good starting point. More often these days computer packages are used to solve electromagnetic field problems with complex boundary conditions.

Engineers generally prefer to work with circuit theory than with field theory. This is typified by the use of 'j notation' to extend the methods of analysis of d.c. circuits to a.c. problems. Microwave engineers normally work with transmission line equivalent circuits whenever possible. A major concern in this book is with the ways in which these equivalent circuits are developed.

The content of the book may be divided into three parts. In Chapters 1 to 4 the emphasis is on basic properties of electromagnetic waves. Chapters 5 to 10 deal systematically with the applications of the theory to a wide range of components and devices. Many of the applications are in microwave engineering but optical and e.m.c. topics are included wherever appropriate. Finally, Chapters 11 and 12 discuss microwave and e.m.c. measuring techniques and provide an overview of the applications of electromagnetic waves in a variety of systems. The aim throughout has been to provide the reader with the basic knowledge which will make the professional literature of the subject accessible. An extensive list of references is provided for this purpose. A small number of exercises are grouped at the end of each chapter (except Chapter 9 where they seemed inappropriate). These are mostly very straightforward. Their purpose is to help the reader to understand the main points in the text, to give confidence in handling the ideas and to give a feel for the numbers involved.

This book carries on the development of the subject from the point I reached in *Electromagnetism for Electronic Engineers* (1986). Those who have found that book helpful will find that the approach in this one is familiar. I hope that it will be found useful not only by students but also by those who discover later in their careers a need for a knowledge of electromagnetic theory.

I am indebted to a number of people who have influenced my own understanding of electromagnetism. My father taught me the value of 'thinking from first principles'. As an undergraduate I relied heavily on Bleaney and Bleaney (1976) which is a model of elegant simplicity in its treatment of the subject. I hope I may have achieved for engineers what that book did for students of physics. My present head of department, Colin Hannaford, was the first to draw my attention to the value of equivalent circuit methods. Finally, Schelkunoff (1943) introduced me to the idea of using transmission-line methods for electromagnetic wave problems.

In the writing of this book I have been heavily indebted to a number of people. Dr L.G. Ripley of the University of Sussex kindly commented on the manuscript and made many helpful suggestions. Dominic Recaldin and his staff at the publishers were patient beyond belief with an author who had a chronic inability to meet deadlines. Most of all I must thank my wife and family who had to put up with my frequent and lengthy disappearances into my study.

R.G. Carter
1989

1 Electromagnetic waves

1.1 INTRODUCTION

This book is about electromagnetic waves, the spectrum of radiation which ranges from the longest radio waves through the infrared and optical regions and on to hard X-rays and gamma rays. Figure 1.1 shows the electromagnetic spectrum and some of its uses. Engineers make use of every part of this vast range of frequencies in information systems of all kinds. Although it is sometimes possible to work with the systems without knowledge of the underlying physical principles there are occasions when this ignorance is a handicap. The whole subject is based on just four physical laws and the consequences of their application to problems with different boundary conditions. Many of these problems can be studied without the use of advanced mathematical methods. Indeed the use of those methods can hinder the growth of the physical understanding which really solves problems. The aim of this book is three-fold: to help the reader to gain a physical understanding of problems involving electromagnetic waves, to relate that understanding particularly to modern problems, and to provide a route into the professional literature of the subject.

Engineers first became aware of electromagnetic waves in the middle of the nineteenth century with the development of the electric telegraph. This was understood, however, in terms of circuit theory rather than electromagnetic field theory. At that time optics was regarded as a separate branch of physics. The work of James Clerk Maxwell, published in his Treatise on Electricity and Magnetism in 1873, provided for the first time a field theory of electromagnetic waves and evidence that light is also an electromagnetic phenomenon. The subsequent exploitation of that theory in radio, radar, television, satellite communications and coherent optics has produced a transformation of human life which is still going on. Maxwell's equations are therefore the starting point of any discussion of these subjects and also of the unwanted electromagnetic coupling which is of increasing concern to electronic engineers.

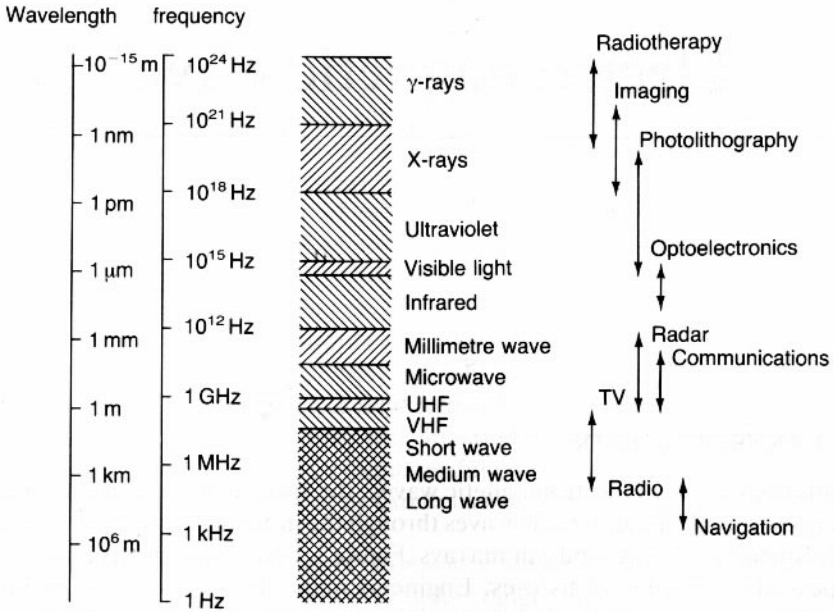

Fig. 1.1 Chart of the electromagnetic spectrum showing some of the uses to which electromagnetic waves are put.

1.2 MAXWELL'S EQUATIONS

The equations which are now known as Maxwell's equations are actually a summary of the four basic laws of electromagnetism (Carter, 1986 p. 12). These are listed below.

1. Gauss' theorem of electrostatics states that the flux of the electric flux density $\boldsymbol{D}$, sometimes known as the electric displacement, out of a closed surface is equal to the total free charge enclosed.

$$\oint\!\!\!\oint \boldsymbol{D}\cdot \mathrm{d}\boldsymbol{A} = \iiint \varrho \,\mathrm{d}v, \tag{1.1}$$

where ϱ is the charge density and $\mathrm{d}\boldsymbol{A}$ and $\mathrm{d}v$ are elements of the area of the surface and of its volume.

2. Gauss' theorem of magnetostatics states that the flux of the magnetic flux density $\boldsymbol{B}$ out of a closed surface is zero (because magnetic monopoles do not exist).

$$\oint\!\!\!\oint \boldsymbol{B}\cdot \mathrm{d}\boldsymbol{A} = 0. \tag{1.2}$$

3. The magnetic circuit law as modified by Maxwell to include the displacement current (sometimes known as Ampère's law) states that the line integral of the magnetic field vector $\boldsymbol{H}$ around a closed path is equal to the total current flux (conduction plus displacement current) through a surface bounded by that path.

$$\oint \boldsymbol{H}\cdot \mathrm{d}\boldsymbol{l} = \iint \left(\boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}\right)\cdot \mathrm{d}\boldsymbol{A}, \tag{1.3}$$

where $\boldsymbol{J}$ is the conduction current density.

4. Faraday's law of electromagnetic induction states that the line integral of the electric field vector $\boldsymbol{E}$ around a closed path is equal to the rate of change of the magnetic flux through a surface bounded by that path

$$\oint \boldsymbol{E}\cdot \mathrm{d}\boldsymbol{l} = -\iint \frac{\partial \boldsymbol{B}}{\partial t}\cdot \mathrm{d}\boldsymbol{A}. \tag{1.4}$$

These are the integral forms of the equations. If the notation is found intimidating it is helpful to remember that it is just a way of writing the usual statements of the laws of electromagnetism in the shorthand notation of mathematics. The laws may also be written in equivalent, differential, forms:

1. Gauss' theorem of electrostatics

$$\nabla\cdot \boldsymbol{D} = \varrho; \tag{1.5}$$

2. Gauss' theorem of magnetostatics

$$\nabla\cdot \boldsymbol{B} = 0; \tag{1.6}$$

3. The magnetic circuit law

$$\nabla \wedge \boldsymbol{H} = \boldsymbol{J} + \frac{\partial \boldsymbol{D}}{\partial t}; \text{ and} \tag{1.7}$$

4. Faraday's law of electromagnetic induction

$$\nabla \wedge \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}. \tag{1.8}$$

The notation of vector calculus used above may, again, be rather intimidating to those who are not mathematically minded. The important thing to remember is that these expressions can be given meanings in a variety of systems of coordinates. Appendix B summarizes these interpretations in Cartesian, and cylindrical and spherical polar coordinates, those being the ones most commonly used by engineers. For the greater part of this book only rectangular Cartesian coordinates are required.

Maxwell's equations are believed to be expressions of basic physical laws. In order to make use of them we also require another set of equations which summarize experimental information about the properties of materials (Carter, 1986 p. 130).

1. Ohm's law

$$J = \sigma E, \tag{1.9}$$

where σ is the conductivity of the material. In this book σ is used rather than its reciprocal, the resistivity ϱ, to avoid confusion with the use of that symbol for charge density.

2. Dielectric materials

$$D = \varepsilon E, \tag{1.10}$$

where ε is the permittivity of the material which can also be written as

$$\varepsilon = \varepsilon_0 \varepsilon_r, \tag{1.11}$$

where ε_0 is the primary electric constant and ε_r is the relative permittivity of the material.

3. Magnetic materials

$$B = \mu H, \tag{1.12}$$

where μ is the permeability of the material which can also be written as

$$\mu = \mu_0 \mu_r, \tag{1.13}$$

where μ_0 is the primary magnetic constant and μ_r is the relative permeability of the material. For some materials either or both of ε_r and μ_r may be complex indicating a phase difference between D and E or B and H for alternating fields.

It is important to bear in mind that these equations are useful approximations of experimental results. They assume that the material properties are constants. While this is a satisfactory assumption for many conducting and dielectric materials it is only a crude approximation for ferromagnetic and ferroelectric materials. In some cases the material properties cannot be regarded as scalar quantities. The vectors which are related to each other by an equation are then not parallel to each other. These aspects of the subject are beyond the scope of this book. Here we shall be assuming that all the properties of materials can be described by scalar constants. (See Dekker, 1959, for further information)

The final equation which will be needed summarizes the law of conservation of charge:

The net current flow out of a closed surface is equal to the rate of decrease of the enclosed charge.

In mathematics this is expressed by the continuity equation which can be written in both integral and differential forms

$$\oiint J \cdot dA = -\frac{\partial}{\partial t} \iiint \varrho \, dv \tag{1.14}$$

$$\nabla \cdot J = -\frac{\partial \varrho}{\partial t}. \tag{1.15}$$

The remainder of this chapter explores the possibility of plane wave solutions to Maxwell's equations in a variety of media.

1.3 ELECTROMAGNETIC WAVES IN NON-CONDUCTING MEDIA

The field theory of a lossless two-wire transmission line assumes that the electric and magnetic fields are perpendicular to each other and to the direction of the line (Carter, 1986, Ch. 7). The results of this theory are consistent with experiment and with the parallel approach using circuit theory summarized in Appendix A. A wave of this kind is called a transverse electric and magnetic (TEM) wave. It is natural to enquire whether this result can be generalized to other situations.

The exploration starts from equations (1.7) and (1.8). By restricting attention to non-conducting materials (1.7) can be simplified to give

$$\nabla \wedge H = \frac{\partial D}{\partial t}. \tag{1.16}$$

To simplify matters still further we assume that, working in rectangular Cartesian coordinates, E only has a component in the x direction and H only in the y direction. If we also assume that any wave propagates in the z direction it follows that the two field vectors vary only with z and t.

In rectangular Cartesian coordinates, the curl of H can be written in the form of a determinant (see Appendix B)

$$\nabla \wedge H = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ H_x & H_y & H_z \end{vmatrix}, \tag{1.17}$$

where $\hat{x}$, $\hat{y}$ and $\hat{z}$ are unit vectors in the x, y and z directions, respectively. Evaluation of the determinant with the assumed direction of H yields

$$\nabla \wedge H = -\hat{x}\, \partial H_y/\partial z \tag{1.18}$$

since $\partial H_y/\partial x$ and $\partial H_y/\partial y$ are both zero. Notice that this vector is in the x direction. Equation (1.16) therefore becomes

$$\frac{\partial H_y}{\partial z} = -\varepsilon \frac{\partial E_x}{\partial t}. \tag{1.19}$$

This is a scalar equation because the vectors on either side are both in the x direction.

In the same way equation (1.8) leads to

$$\frac{\partial E_x}{\partial z} = -\mu \frac{\partial H_y}{\partial t}. \tag{1.20}$$

These equations are remarkably similar in form to the telegrapher's equations derived from the circuit theory of transmission lines (Carter, 1986, p. 108)

$$\frac{\partial I}{\partial x} = -C \frac{\partial V}{\partial t} \tag{1.21}$$

$$\frac{\partial V}{\partial x} = -L \frac{\partial I}{\partial t}. \tag{1.22}$$

The resemblance is even more striking when the dimensions of the quantities are recalled: $\boldsymbol{H}$ in amps per metre, $\boldsymbol{E}$ in volts per metre, ε in farads per metre and μ in henries per metre.

Differentiating (1.19) with respect to t and (1.20) with respect to z gives

$$\frac{\partial^2 H_y}{\partial z \partial t} = -\varepsilon \frac{\partial^2 E_x}{\partial t^2} \tag{1.23}$$

$$\frac{\partial^2 E_x}{\partial z^2} = -\mu \frac{\partial^2 H_y}{\partial z \partial t}, \tag{1.24}$$

whence

$$\frac{\partial^2 E_x}{\partial z^2} = \varepsilon\mu \frac{\partial^2 E_x}{\partial t^2}. \tag{1.25}$$

This is the standard form of the wave equation. For sinusoidal waves the solution can be written (Carter, 1986, p. 110)

$$E_x = E_+ \exp \mathrm{j}(\omega t - kz) + E_- \exp \mathrm{j}(\omega t + kz), \tag{1.26}$$

where E_+ and E_- are the amplitudes of waves travelling in the positive and negative z directions, respectively. It should be remembered that the use of complex notation is a convenient way of simplifying the mathematical manipulations and that here, as in the 'j notation' used in circuit theory, it is the real part of every expression which is taken to have physical significance.

The phase velocity of the waves is given by

$$v_p = \frac{\omega}{k} = \frac{1}{\surd(\varepsilon\mu)} \tag{1.27}$$

(c.f. the relationship $v_p = 1/\surd(LC)$ for a transmission line).

For the special case of waves travelling in free space, the numerical value of the phase velocity calculated from experimental values of ε_0 and μ_0 agrees with the experimental measurements of the velocity of light within the limits of experimental error. It is easy to show that H_y also satisfies an equation like (1.25).

The relationship between E_x and H_y can be found by substituting the general solution (1.26) and the equivalent expression for H_y into equation (1.19). Then for waves travelling in the positive direction

$$\mathrm{j}kH_y = \mathrm{j}\omega\varepsilon E_x \tag{1.28}$$

or

$$\frac{E_x}{H_y} = \sqrt{\left(\frac{\mu}{\varepsilon}\right)} = Z_w. \tag{1.29}$$

This quantity has the dimensions of impedance and it is referred to as the wave impedance. In free space it has the numerical value 377 Ω. One important deduction can be made from (1.29) namely that E_x and H_y are in phase with each other. It can be shown that the wave impedance of the wave in a transmission line also satisfies (1.29). (Carter, 1986). For waves travelling in the $-z$ direction k is negative and a minus sign appears in (1.28). The relationship between E_x and H_y in a plane electromagnetic wave is usually represented by the diagram shown in Fig. 1.2. This diagram shows the field vectors varying sinusoidally in space with the whole pattern moving in the z direction at the phase velocity of the wave. This diagram is a little misleading because it does not give the impression that $\boldsymbol{E}_y$ and $\boldsymbol{H}_y$ are constant over any plane perpendicular to the z axis. At this point it may be asked how it is possible for the field lines to lie exactly in a plane and,

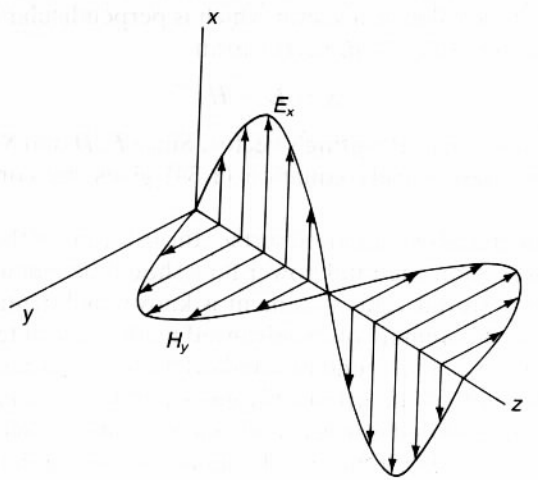

Fig. 1.2 Relationship between the electric and magnetic field vectors in a plane electromagnetic wave.

consequently, never end. The answer to this is that, in practice, a plane wave is created by launching a spherical wave. At a large distance from the source any small part of the wave front is effectively a plane wave.

1.4 ENERGY FLOW IN AN ELECTROMAGNETIC WAVE

When a wave $V \exp j(\omega t - kz)$ propagates on a transmission line of characteristic impedance Z_0 the instantaneous flow of power along the line given by

$$P = VI = V^2/Z_0 = Z_0 I^2. \tag{1.30}$$

It is plausible to suppose that, in the same way, an electromagnetic wave propagating in free space also carries power. Now in a region where the electric and magnetic field strengths are $\boldsymbol{E}$ and $\boldsymbol{H}$ the stored energy density is

$$W = \tfrac{1}{2}\varepsilon|E|^2 + \tfrac{1}{2}\mu|H|^2. \tag{1.31}$$

By making use of the relationship (1.29) between $\boldsymbol{E}$ and $\boldsymbol{H}$ we find that the peak energy density in an electromagnetic wave is

$$W = \surd(\varepsilon\mu)\ |\boldsymbol{E}|\,|\boldsymbol{H}|. \tag{1.32}$$

The peak power flow is obtained by multiplying this expression by the velocity of the wave from (1.27) to give

$$S = |\boldsymbol{E}|\,|\boldsymbol{H}|. \tag{1.33}$$

Strictly speaking the group velocity should be used here in place of the phase velocity (see p. 40) but for waves propagating in uniform dielectric media they are identical.

Because the power flow is a vector which is perpendicular to both $\boldsymbol{E}$ and $\boldsymbol{H}$ it is useful to write (1.33) in vector form

$$\boldsymbol{S} = \boldsymbol{E} \wedge \boldsymbol{H}. \tag{1.34}$$

The vector $\boldsymbol{S}$ is known as Poynting's vector. Since $\boldsymbol{E}$, $\boldsymbol{H}$ and $\boldsymbol{S}$ are in the x, y and z directions respectively equation (1.34) gives the correct direction for $\boldsymbol{S}$.

When S is integrated over a closed surface the principle of the conservation of energy requires that the total power flow should be equal to the rate of change of energy stored. This statement is known as Poynting's theorem. The proof of the theorem involves advanced mathematical techniques so it is not given here; it can be found in standard texts (e.g. Ramo *et al.*, 1965). It should be noted that the proof only gives a physical significance to the flux of $\boldsymbol{S}$ out of a closed surface and not to $\boldsymbol{S}$ itself. Where energy is dissipated as heat within the surface that must also be taken into account in applying the principle of conservation of energy. Poynting's theorem can be regarded as a generalization of the ideas of the flow of energy in electric

circuits which associates the flow with the fields rather than directly with the currents and voltages.

The Poynting vector is not the only possible expression of the power flow in an electromagnetic wave. An alternative, the Slepian vector, is discussed by Carter (1967). The two approaches correspond to rather different physical pictures of the way in which energy is transmitted and dissipated. In practice the Poynting vector is the one generally used.

Very often the electromagnetic power flow is that in a sinusoidal electromagnetic wave. In that case the time average Poynting vector is

$$\langle \boldsymbol{S} \rangle = \tfrac{1}{2}\boldsymbol{E} \wedge \boldsymbol{H}^* \tag{1.35}$$

where $\boldsymbol{E}$ and $\boldsymbol{H}$ are the complex wave vectors, the asterisk indicates the complex conjugate, and the factor of $\frac{1}{2}$ is a consequence of averaging the power flow over a full cycle of the wave. This expression is closely analogous to the usual expression for the power flow in an electric circuit

$$P = \tfrac{1}{2}VI^* \tag{1.36}$$

demonstrating again that the equations of circuit theory are special cases of the general laws of electromagnetism which apply when the currents are constrained to flow in wires and the components of the circuit can be regarded as lumped.

1.5 ELECTROMAGNETIC WAVES IN CONDUCTING MATERIALS

If the wave propagates in a conducting material then it is necessary to include the conduction current density $\boldsymbol{J}$ in the equations. If we again assume that the electric and magnetic fields are in the x and y directions respectively and that they vary as $\exp j(\omega t - kz)$ then (1.7) becomes

$$jkH_y = (\sigma + j\omega\varepsilon)E_x. \tag{1.37}$$

A good conductor may be defined as a material in which the conduction current is much greater than the displacement current, that is $\sigma \gg j\omega\varepsilon$ in (1.37) so that, approximately

$$jkH_y = \sigma E_x. \tag{1.38}$$

For copper $\sigma = 5.7 \times 10^7\,\mathrm{S\,m^{-1}}$ and $\varepsilon = 10^{-11}\,\mathrm{F\,m^{-1}}$ so that the approximation is valid up to frequencies around $10^{16}\,\mathrm{Hz}$. Similarly, from (1.8) we get

$$jkE_x = j\omega\mu H_y \tag{1.39}$$

so that

$$k^2 = -j\omega\sigma\mu$$
$$k = \pm\surd{-j}\ \surd(\omega\sigma\mu). \tag{1.40}$$

The square root of $-j$ can be found by noting that

$$-j = \exp(-j\pi/2) \tag{1.41}$$

so that

$$\begin{aligned}\sqrt{-j} &= \exp(-j\pi/4)\\ &= \cos(-\pi/4) + j\sin(-\pi/4)\\ &= \frac{1}{\sqrt{2}}(1-j)\end{aligned} \tag{1.42}$$

and

$$\begin{aligned}k &= \pm(1-j)\sqrt{\left(\frac{\omega\sigma\mu}{2}\right)}\\ &= \pm(1-j)/\delta,\end{aligned} \tag{1.43}$$

where $$\delta = \sqrt{(2/\omega\sigma\mu)}. \tag{1.44}$$

Substituting for k in the expressions for the fields shows that E and H vary (for waves travelling in the $+z$ direction) as

$$\exp j(\omega t - z/\delta)\exp(-z/\delta).$$

Thus a wave propagates with an exponentially decreasing amplitude. The decay constant δ is known as the skin depth for reasons which will become apparent in Chapter 4. The skin depth varies with frequency and with the properties of the material. For copper we find

at 1 kHz	$\delta = 2$ mm
at 1 MHz	$\delta = 67\,\mu$m
at 1 GHz	$\delta = 2\,\mu$m.

The wave impedance is given by

$$\begin{aligned}Z_w &= E_x/H_y = jk/\sigma\\ &= (1+j)/\sigma\delta\end{aligned} \tag{1.45}$$

so that the electric and magnetic fields are not in phase with each other but the electric field leads the magnetic field by 45°.

The power flow in the wave is given by the real part of the complex Poynting vector

$$\begin{aligned}|\boldsymbol{S}| &= \mathrm{Re}|\tfrac{1}{2}\boldsymbol{E}\wedge\boldsymbol{H}^*|\\ &= \mathrm{Re}\left[\tfrac{1}{2}|\boldsymbol{E}|^2\left(\frac{\sigma\delta}{1+j}\right)\right]\\ &= \mathrm{Re}[\tfrac{1}{4}|\boldsymbol{E}|^2\sigma\delta(1-j)]\\ &= \tfrac{1}{4}|\boldsymbol{E}|^2\sigma\delta.\end{aligned} \tag{1.46}$$

In some materials, mainly lossy dielectrics, the conduction and displacement currents are of comparable magnitudes so that the solution for k must be derived from (1.37) and (1.39). An added complication is that, because of the effects of vibration and rotation of molecules, $\boldsymbol{D}$ and $\boldsymbol{E}$ are not necessarily in phase with each other. This is allowed for by writing

$$\varepsilon = \varepsilon' - j\varepsilon''$$

so that (1.37) becomes

$$\begin{aligned}jkH_y &= [(\sigma + \omega\varepsilon'') + j\omega\varepsilon']E_x\\ &= (\sigma' + j\omega\varepsilon')E_x.\end{aligned} \tag{1.47}$$

The relationship between the conduction and displacement currents is illustrated in Fig. 1.3. The angle between the total current and the displacement current is given by

$$\tan\delta = \sigma'/\omega\varepsilon'. \tag{1.48}$$

This quantity is known as the loss tangent of the material. Note carefully that δ here is not the skin depth referred to above. Equation (1.47) can therefore be written

$$jkH_y = (1 - j\tan\delta)j\omega\varepsilon' E_x \tag{1.49}$$

so that

$$\begin{aligned}k^2 &= (1 - j\tan\delta)\omega^2\varepsilon'\mu\\ k &= \sqrt{(1 - j\tan\delta)}\omega\sqrt{(\varepsilon'\mu)}\end{aligned} \tag{1.50}$$

for many materials the loss tangent is small so that

$$k \simeq (1 - \tfrac{1}{2}j\tan\delta)\omega\sqrt{(\varepsilon'\mu)}. \tag{1.51}$$

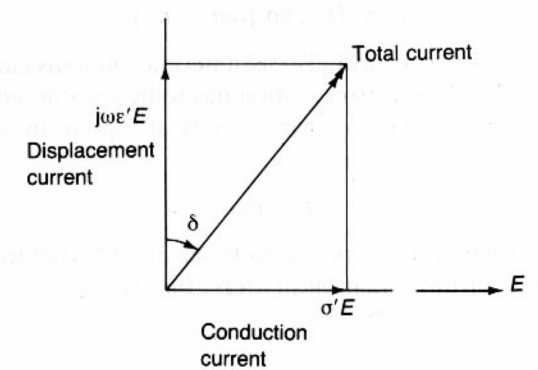

Fig. 1.3 Phasor diagram of the conduction and displacement currents in a lossy dielectric material.

Substituting this expression into the solution assumed for the propagation of the waves we find

$$\exp j\omega[t - \sqrt{(\varepsilon'\mu)}z] \exp -[\tfrac{1}{2}\omega \cdot \sqrt{(\varepsilon'\mu)} \tan \delta]z \tag{1.52}$$

so that the waves decay as they propagate through the material. Typical materials include plastics, ceramics and many organic materials. This part of the theory finds application in the use of electromagnetic energy in microwave ovens. For example, steak has $\varepsilon' = 40\varepsilon_0$ and $\tan \delta = 0.3$ at 3 GHz. Substituting these figures into the expression above shows that the fields fall off by a factor $1/e$ in 17 mm.

1.6 PROPAGATION OF WAVES IN PLASMAS

So far it has been assumed that the conduction charges in a conductor are able to respond instantly to the field of an electromagnetic wave. This assumption is not always valid. A particular case is that of an ionized gas in which there are two species of mobile charge carriers, free electrons and the very much more massive positive ions. It is convenient to speak of such a gas as a plasma although some would restrict the use of that term to situations where the ions are completely stripped of their electrons. The theory of the interaction between electromagnetic waves and plasmas is important for understanding the propagation of waves in the ionosphere, in electron devices and in experiments in thermonuclear fusion. Some of these cases are rather difficult because they involve the random thermal motions of the charge carriers and collisions between them. Here the basic ideas are illustrated by considering a cold, collisionless plasma.

We assume, as before, that the wave is a pure TEM wave propagating in the z direction so that

$$E_x = E_0 \exp j(\omega t - kz)$$

and

$$H_y = H_0 \exp j(\omega t - kz). \tag{1.53}$$

To make things simpler we will also assume that the ions can be regarded as fixed so that only the electron motion has to be considered. The current density, charge density and electron velocity at a point in the plasma are related by

$$\boldsymbol{J} = \varrho \boldsymbol{v}. \tag{1.54}$$

The plasma as a whole is supposed to be electrically neutral so only the time-varying part of this equation matters, that is

$$J_1 = \varrho_0 v_1 + \varrho_1 v_1. \tag{1.55}$$

Provided that the signal levels are small $\varrho_1 \ll \varrho_0$ and the second term is negligible compared with the first so that

$$J_1 = \varrho_0 v_1. \tag{1.56}$$

It can be shown that the magnetic forces on the electrons are much smaller than the electric forces so that, effectively,

$$m\ddot{x} = qE_x$$

or

$$j\omega m v_1 = qE_x. \tag{1.57}$$

Eliminating v_1 between (1.56) and (1.57) gives

$$J_1 = -j\frac{\eta\varrho_0}{\omega}E_x, \tag{1.58}$$

where $\eta = q/m$ is the charge to mass ratio of the electron. When the current and the magnetic field from (1.53) are substituted into (1.7) we get

$$jkH_y = j\omega\varepsilon_0\left(1 - \frac{\omega_p^2}{\omega^2}\right)E_x, \tag{1.59}$$

where

$$\omega_p^2 = \eta\varrho_0/\varepsilon_0. \tag{1.60}$$

This frequency is known as the plasma frequency.

A second relationship between the two field vectors is given by (1.39) and they may then be eliminated to give the propagation constant of the wave

$$k = \omega\sqrt{(\varepsilon_0\mu_0)}\left(1 - \frac{\omega_p^2}{\omega^2}\right)^{\frac{1}{2}}. \tag{1.61}$$

From this it is clear that the propagation constant has real values only when the signal frequency is greater than the plasma frequency. The plasma then behaves as a dielectric medium whose permittivity, from comparison with (1.27), is

$$\varepsilon = \varepsilon_0\left(1 - \frac{\omega_p^2}{\omega^2}\right). \tag{1.62}$$

At frequencies below the plasma frequency k is pure imaginary so that the waves decay as

$$\exp\left[-\left(\frac{\omega_p^2}{\omega^2} - 1\right)^{\frac{1}{2}} k_0 z\right], \tag{1.63}$$

where k_0 is the free-space propagation constant at frequency ω. Further light is shed on this behaviour by consideration of the wave impedance and of the relationship between the conduction and displacement current densities. The wave impedance is, from (1.59) and (1.61),

$$Z_w = \sqrt{\left(\frac{\mu_0}{\varepsilon_0}\right)}\left(1 - \frac{\omega_p^2}{\omega^2}\right)^{\frac{1}{2}}. \tag{1.64}$$

Above the plasma frequency this is real so that $\boldsymbol{E}$ and $\boldsymbol{H}$ are in phase with each other. Below the plasma frequency it is pure imaginary so that $\boldsymbol{E}$ and $\boldsymbol{H}$ are in phase quadrature and there is no net flow of power. The displacement current is

$$J_d = \frac{\partial \boldsymbol{D}}{\partial t} = j\omega\varepsilon_0 E_x, \qquad (1.65)$$

so that, from (1.58), the conduction current is

$$J_1 = -\frac{\omega_p^2}{\omega^2} J_d. \qquad (1.66)$$

Thus below the plasma frequency the electrons can follow the wave, the conduction current dominates and the plasma behaves as a conductor. Above the plasma frequency they cannot do so, the diplacement current dominates and the plasma behaves as a dielectric.

Similar behaviour is seen in some dielectric materials which contain molecules having an electric dipole moment. The permittivities of these materials vary with frequency according to whether the molecules can rotate to follow the changing electric field or not. For further information consult the book by Bleaney and Bleaney (1976).

1.7 POLARIZATION OF WAVES

So far we have only discussed waves which have their electric and magnetic field vectors in the x and y directions, respectively. It is obvious that this is not the only possible orientation of the field vectors for a wave propagating in the positive z direction. The orientation of the electric field vector is referred to by electronic engineers as the plane of polarization of the wave. Somewhat confusingly the convention adopted in optics is to define the plane of polarization as the direction of the magnetic field vector. In this book the first convention will be used throughout.

Any general direction of polarization can be considered as a superposition of two waves having the same phase as each other and polarized in the x and y directions as shown in Fig. 1.4. Such a wave is known as a plane-polarized wave. We shall see later that waves having different polarizations behave differently when they pass through certain media and when they are reflected from the interface between two dielectric materials. These properties have important practical consequences.

Now consider the slightly more complicated case where the two components of a wave are equal in amplitude but have phases which differ from each other by 90°. At a particular plane perpendicular to the z axis the electric field of the wave is

$$\boldsymbol{E} = \hat{\boldsymbol{x}}E_0 \cos \omega t + \hat{\boldsymbol{y}}E_0 \sin \omega t \qquad (1.67)$$

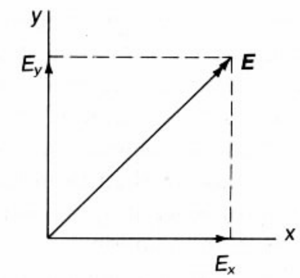

Fig. 1.4 Combination of waves polarized in the x and y directions.

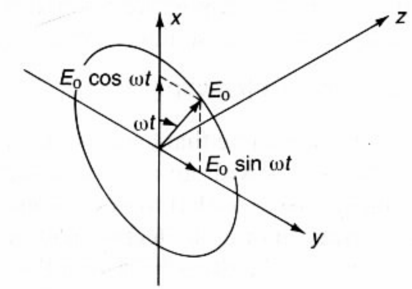

Fig. 1.5 Relationship between the x and y components of the electric field in a circularly polarized wave.

so that the tip of the electric field vector is rotating around a circle with angular velocity ω as shown in Fig. 1.5. The wave is then said to be circularly polarized. Moreover, because the direction of rotation of the electric field vector with time is in the right hand corkscrew sense with respect to the z axis, it is positive circularly polarized. Evidently, if the phase difference between the x and y components had been made $-90°$ the sense of rotation would have been reversed and the resulting wave would have been negative circularly polarized.

In the most general case of all the amplitudes of the two components may differ from each other and their phases be other than in quadrature. The tip of the electric field vector then traces out an ellipse and the wave is said to be elliptically polarized. As this increase in generality introduces no new principles it will not be pursued further here and the reader is referred to more advanced texts for the details (e.g. Jordan and Balmain, 1968; Longhurst, 1973)

1.8 PROPAGATION IN GYROMAGNETIC MEDIA

Some media have the property that when they are placed in a steady magnetic field the propagation constants for positive and negative circularly polarized waves differ from each other. Examples are ionized gasses and ferrites. They are known collectively as gyromagnetic materials for reasons which will become apparent. Ferrites are made by sintering mixtures of oxides of iron and of metals such as nickel or magnanese (Baden-Fuller, 1987). They combine ferromagnetic properties with high electrical resistivity and have been developed because of their usefulness at high frequencies. The mathematical treatment of wave propagation in ferrites is rather involved so the case of propagation in an ionized gas in a magnetic field will be used here to illustrate how the gyromagnetic properties arise. Wave propagation in ferrites is discussed in Chapter 8.

Consider, then, an ionized gas which is in a steady magnetic field B_0 directed parallel to the z axis. Let a positive circular polarized wave pass through in the z direction. The electric field of the wave is then

$$\boldsymbol{E} = [\hat{x} - \mathrm{j}\hat{y}]E_0 \exp \mathrm{j}(\omega t - kz). \tag{1.68}$$

To simplify the derivation we assume that each electron moves in a circular orbit perpendicular to the z axis with an angular velocity ω. The ions are assumed to be sufficiently massive so that their motion can be ignored. The electrons experience the rotating radial electric field of the wave shown in Fig. 1.5 and the condition for a steady orbit is that the radial forces should balance. That is

$$-qE_0 - qr\omega B_0 + mr\omega^2 = 0, \tag{1.69}$$

where q is the magnitude of the charge on an electron. The three terms represent the electric force, the magnetic force produced by the motion of the electron through the magnetic field, and the centrifugal force. Equation (1.69) can be rearranged to give the radius of the stable orbit

$$r = \frac{\eta E_0}{\omega(\omega - \omega_c)}, \tag{1.70}$$

where η is the charge to mass ratio of the electron and

$$\omega_c = \eta B_0 \tag{1.71}$$

is the cyclotron frequency, that is, the angular velocity of the electron for a stable orbit when the electric field is zero. Equation (1.70) shows that the radius of the orbit is greatest when the signal frequency is equal to the cyclotron frequency. It would not become infinite in practice because of the effects of collisions. When $\omega > \omega_c$ r is positive so the electron moves in phase with the applied field. When $\omega < \omega_c$ r is negative, implying that the motion and the field are in antiphase.

Figure 1.6 shows the Cartesian components of the velocity of the electron. If there are N electrons per unit volume then the x component of the current density is

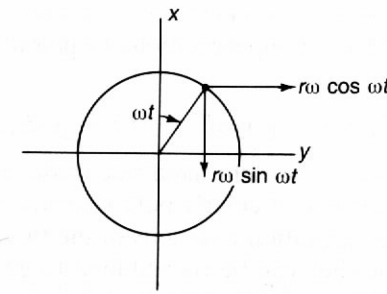

Fig. 1.6 Cartesian components of the velocity of an electron moving in a circular orbit.

$$\begin{aligned} J_x &= -Nq\dot{x} \\ &= -\mathrm{j}Nq\omega r \exp \mathrm{j}(\omega t - kz). \end{aligned} \tag{1.72}$$

Then, from (1.7), we get

$$\mathrm{j}kH_y = J_x + \mathrm{j}\omega\varepsilon_0 E_x \tag{1.73}$$

or

$$kH_y = \omega\varepsilon_0 E_x\left[1 - \frac{\omega_p^2}{\omega(\omega - \omega_c)}\right]. \tag{1.74}$$

Combining this with (1.39) gives the propagation constant of the wave (c.f. eqn (1.61))

$$k_+ = \omega\surd(\varepsilon_0\mu_0)\left[1 - \frac{\omega_p^2}{\omega(\omega - \omega_c)}\right]^{\frac{1}{2}}, \tag{1.75}$$

where the subscript + indicates positive circular polarization. Equation (1.75) shows that at very high frequencies the propagation constant tends to that of free space. This is because the electrons are unable to follow the changes in the field.

The equivalent expression for the negative circularly polarized wave is obtained by setting ω equal to $-\omega$, giving

$$k_- = \omega\surd(\varepsilon_0\mu_0)\left[1 - \frac{\omega_p^2}{\omega(\omega + \omega_c)}\right]^{\frac{1}{2}}. \tag{1.76}$$

Thus the positively and negatively polarized waves have different propagation constants. The physical explanation for this effect is revealed when equation (1.69) is examined. For positive rotation the magnetic force is inwards whilst for negative rotation it is outwards. Thus for positive rotation the electric force required to produce equilibrium can be either inwards or outwards, whereas for negative rotation it must always be inwards.

This effect has important practical consequences. A plane polarized wave can be considered as the superposition of a pair of circularly polarized waves by writing

$$\hat{x}E_x = \tfrac{1}{2}(\hat{x} + \mathrm{j}\hat{y})E_x + \tfrac{1}{2}(\hat{x} - \mathrm{j}\hat{y})E_x, \qquad (1.77)$$

where the two brackets on the right hand side of the equation represent a pair of positive and negative circularly polarized waves.

In most media the propagation constants of the two waves are the same and at any other plane they can be recombined to give a plane polarized wave with the same plane of polarization as before. In a gyromagnetic medium, however, the propagation constants of the waves differ, as we have seen, so that the phase relationship between them changes as they propagate. When they are recombined the effect is to produce a plane polarized wave whose plane of polarization has been rotated about the z axis relative to the initial polarization. This effect is known as Faraday rotation.

In the ionosphere the Earth's atmosphere is ionized by cosmic rays to produce a plasma which is influenced by the Earth's magnetic field. Thus the ionosophere is a gyromagnetic medium of the kind discussed above. Radio signals transmitted from satellites experience a rotation in their planes of polarization as they pass through the ionosphere and the extent of the rotation varies with time because of variations in the density of free electrons. If plane waves were used for the transmissions there would be difficulties in receiving them because of this effect. The solution is to transmit circularly polarized waves so that the Faraday rotation appears only as a phase shift in the signal received.

Faraday rotation also occurs when a wave passes through a ferrite material which is in a steady magnetic field. This is put to use in the microwave devices known as circulators and isolators which are discussed further in Chapter 8.

1.9 BOUNDARY CONDITIONS

The study of waves in uniform, infinite media is of limited interest. Practical problems usually involve more than one medium so that the behaviour of the waves at the interfaces is very important. The starting point for discussing this subject is the statement of the boundary conditions which apply to the electric and magnetic fields at a boundary at which there are no surface charges or currents.

1. The tangential component of the electric field $\boldsymbol{E}$ is continuous.
2. The normal component of the electric flux density $\boldsymbol{D}$ is continuous.
3. The tangential conponent of the magnetic field $\boldsymbol{H}$ is continuous.
4. The normal component of the magnetic flux density $\boldsymbol{B}$ is continuous.

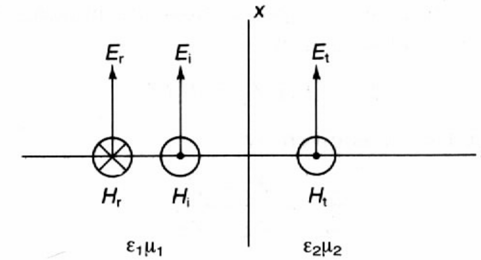

Fig. 1.7 Incident, reflected and transmitted wave fields for an electromagnetic wave incident normally on a dielectric boundary.

The proofs of these conditions can be found in books on elementary electromagnetism (Carter, 1986, pp. 24 and 66).

The simplest case of interaction between an electromagnetic wave and a boundary arises when the direction of propagation is normal to it. The field vectors are then parallel to the boundary and only boundary conditions 1 and 3 are needed. Figure 1.7 shows this situation with the incident, transmitted and reflected wave fields. The origin of coordinates is taken to lie on the boundary for convenience. The propagation constants in the two materials are

$$k_1 = \omega\surd(\varepsilon_1\mu_1) \quad \text{and} \quad k_2 = \omega\surd(\varepsilon_2\mu_2) \qquad (1.78)$$

and the wave impedances are

$$Z_1 = \sqrt{\left(\frac{\mu_1}{\varepsilon_1}\right)} \quad \text{and} \quad Z_2 = \sqrt{\left(\frac{\mu_2}{\varepsilon_2}\right)}. \qquad (1.79)$$

The three waves are then

$$\begin{aligned} &\boldsymbol{E}_\mathrm{i} \exp \mathrm{j}(\omega t - k_1 z) \\ &\boldsymbol{E}_\mathrm{r} \exp \mathrm{j}(\omega t + k_1 z) \\ &\boldsymbol{E}_\mathrm{t} \exp \mathrm{j}(\omega t - k_2 z). \end{aligned} \qquad (1.80)$$

Boundary condition 1 requires that the electric fields should be the same on both sides of the boundary.

$$\boldsymbol{E}_\mathrm{i} + \boldsymbol{E}_\mathrm{r} = \boldsymbol{E}_\mathrm{t}. \qquad (1.81)$$

Similarly condition 3 yields

$$\boldsymbol{H}_\mathrm{i} - \boldsymbol{H}_\mathrm{r} = \boldsymbol{H}_\mathrm{t}. \qquad (1.82)$$

The equations are analogous to those for the voltage and current at a discontinuity in a transmission line (Appendix A). The magnetic field vectors for the incident and reflected waves must be of opposite sign in order to give the correct directions for the power flow. By making use of the wave impedances (1.82) can be rewritten

$$(E_i - E_r)/Z_1 = E_t/Z_2, \tag{1.83}$$

whence the reflected wave is given by

$$\frac{E_r}{E_i} = \frac{Z_2 - Z_1}{Z_2 + Z_1} \tag{1.84}$$

and the transmitted wave is given by

$$\frac{E_t}{E_i} = \frac{2Z_2}{Z_2 + Z_1}. \tag{1.85}$$

We have seen in the preceding sections that the wave impedance is sometimes a complex or imaginary quantity. When this occurs the ratios of the wave amplitudes are complex indicating reflection and transmission with a change of phase. An important special case arises when the second material is regarded as a perfect conductor. The electric field within it and the wave impedance must then be zero so that

$$\frac{E_r}{E_i} = -1 \quad \text{and} \quad \frac{E_t}{E_i} = 0 \tag{1.86a}$$

that is, the wave is totally reflected at the boundary with the reflected wave in antiphase with the incident wave. The corresponding equations for the magnetic fields are

$$\frac{H_r}{H_i} = 1 \quad \text{and} \quad \frac{H_t}{H_i} = 0. \tag{1.86b}$$

From these it follows that the magnetic field is $2\boldsymbol{H}_i$ just outside the conductor and zero within it. There is, therefore, an apparent violation of boundary condition 3. The explanation is that currents flow in the surface of the conductor to match the fields on either side (see Carter, 1986, p. 66). The surface current density is equal to the tangential component of the magnetic field.

Equations (1.84) and (1.85) bear a striking resemblance to the equations for the transmission and reflection coefficients at a junction between two transmission lines of different impedances (see Carter, 1986, p. 112). This is not surprising when we remember that the waves on the lines can be described either in circuit terms or as TEM field waves. We have already seen that there is a close resemblance between the differential equations governing the two descriptions (eqns (1.19) to (1.21)). The analogy is very important practically because it allows us to make use of transmission line techniques for solving electromagnetic wave problems. This point will be explored further in Chapters 3 and 4.

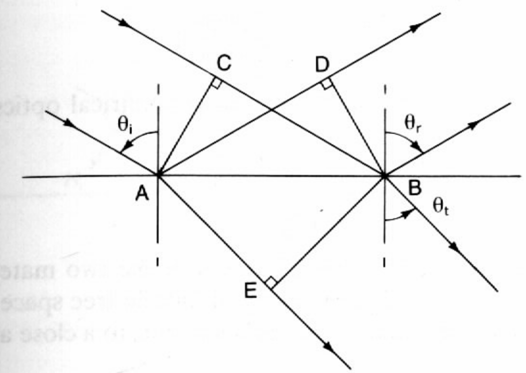

Fig. 1.8 Geometry of the reflection and refraction of waves at a dielectric boundary for oblique incidence.

In general a wave will not be incident normally on a boundary. For other angles of incidence we have to consider separately waves polarized normal to the boundary and waves polarized parallel to the boundary. Any more general case can be regarded as a superposition of these two.

First we shall establish the laws of reflection and refraction by considering Fig. 1.8. The diagram shows general incident, reflected and transmitted waves making angles θ_i, θ_r and θ_t to the normal to the boundary. Points A and B are chosen on the boundary so that the phase difference between them is 360° for the incident wave. Whatever conditions apply at A must also apply at B so that the phase differences for the other two waves must also be 360°. Wavefronts AC, BD and BE are constructed for each wave perpendicular to the directions of propagation. The distances CB, AD and AE are then the distances travelled by the wavefronts in one cycle of the wave, that is

$$CB = AD = \lambda_1$$

and

$$AE = \lambda_2, \tag{1.87}$$

where λ_1 and λ_2 are the wavelengths in the two media. The triangles ABC and ABD are similar triangles and the angles of incidence and reflection are proved to be equal to each other. We can therefore use subscripts 1 and 2 to refer to the angles in the two media.

For the transmitted wave, from triangles ABC and ABE we have

$$AB = \frac{\lambda_1}{\sin\theta_1} = \frac{\lambda_2}{\sin\theta_2} \tag{1.88}$$

or, since

$$\lambda_1 = \frac{2\pi}{\omega\surd(\varepsilon_1\mu_1)} \quad \text{and} \quad \lambda_2 = \frac{2\pi}{\omega\surd(\varepsilon_2\mu_2)} \tag{1.89}$$

$$\frac{\sin\theta_1}{\sin\theta_2} = \sqrt{\left(\frac{\varepsilon_2\mu_2}{\varepsilon_1\mu_1}\right)}. \tag{1.90}$$

This is evidently related to Snell's law of geometrical optics (Longhurst, 1973).

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1}, \tag{1.91}$$

where n_1 and n_2 are the refractive indexes of the two materials. The refractive index is the ratio of the velocity of light in free space to that in the medium. For non-ferromagnetic materials $\mu = \mu_0$ to a close approximation so that $n = \sqrt{\varepsilon_r}$.

Having established these basic relations we can proceed to consider the oblique incidence of a wave on a boundary taking first the case shown in Fig. 1.9 where the plane of polarization is normal to the boundary. Applying boundary conditions 1 and 3 as before yields

$$(E_i + E_r)\cos\theta_1 = E_t\cos\theta_2 \tag{1.92}$$

$$H_i - H_r = H_t. \tag{1.93}$$

Comparison with equations (1.81) and (1.82) shows that we can maintain the correspondence with transmission line theory if we define the normal wave impedance of the waves as

$$Z_{n\perp} = \frac{E\cos\theta}{H} = \sqrt{\left(\frac{\mu}{\varepsilon}\right)}\cos\theta. \tag{1.94}$$

When the electric field vectors are parallel to the boundary as shown in Fig. 1.10 the boundary conditions are

$$E_i + E_r = E_t \tag{1.95}$$

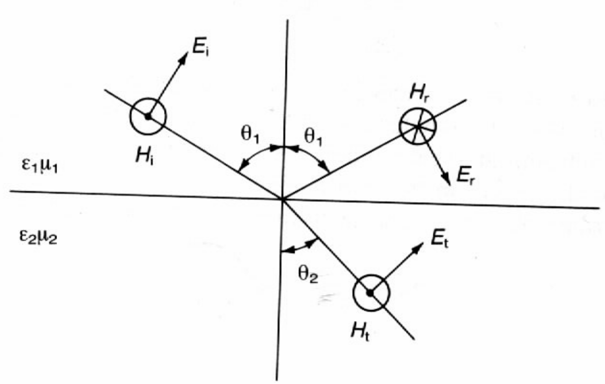

Fig. 1.9 Field vectors for oblique incidence of waves on a dielectric boundary with the electric field normal to the boundary.

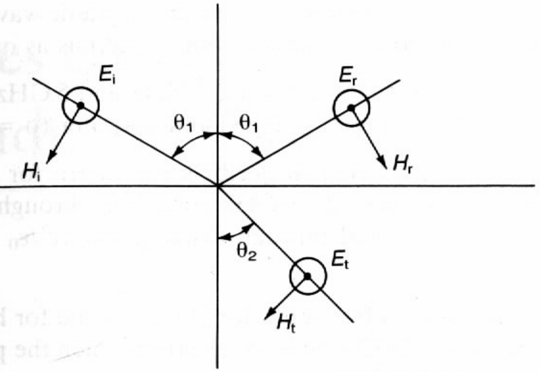

Fig. 1.10 Field vectors for oblique incidence of waves on a dielectric boundary with the electric field parallel to the boundary.

and
$$(H_i - H_r)\cos\theta_1 = H_t\cos\theta_2 \tag{1.96}$$

which can be made equivalent to (1.81) and (1.83) by defining the normal impedances of the waves by

$$Z_{n\parallel} = \frac{E}{H\cos\theta} = \sqrt{\left(\frac{\mu}{\varepsilon}\right)}\frac{1}{\cos\theta}. \tag{1.97}$$

Practical cases involving the interactions between waves and boundaries between different materials will be considered in later chapters.

1.10 CONCLUSION

In this chapter we have considered the propagation of plane electromagnetic waves through different media and seen how it depends upon their physical properties. The different cases considered are not exhaustive but have been chosen to illustrate the principal kinds of phenomena which occur. We have also established the laws of reflection and refraction which apply to waves at the interface between two materials. A close resemblance to transmission line theory has been demonstrated which promises to allow transmission line methods to be applied to problems involving electromagnetic waves. The fundamental concepts considered in this chapter are applied to practical situations in the chapters which follow.

EXERCISES

1.1 Calculate the wave impedances of electromagnetic waves travelling in free space, polystyrene ($\varepsilon_r = 2.7$), alumina ($\varepsilon_r = 8.9$) and Barium strontium titanate ($\varepsilon_r = 10\,000$).

1.2 Calculate the power density in an electromagnetic wave whose electric field strength is $100\,V\,m^{-1}$ in the same materials as question 1.1.

1.3 Calculate the skin depth at 50 Hz, 5 MHz and 5 GHz for silver ($\sigma = 6.1 \times 10^7\,S$), Graphite ($\sigma = 10^5\,S$) and seawater ($\sigma = 4\,S$).

1.4 Calculate the attenuation in decibels per metre for electromagnetic waves at a frequency of 10 GHz travelling through glass ($\varepsilon_r = 4$, $\varepsilon''/\varepsilon' = 21 \times 10^{-4}$) and through fused quartz ($\varepsilon'/\varepsilon_0 = 3.78$, $\varepsilon''/\varepsilon' = 10^{-4}$).

1.5 Calculate the plasma frequency for electrons and for hydrogen molecular ions (mass = 3672 × mass of electron) when the particle densities are $10^{12}\,m^{-3}$ and $10^{16}\,m^{-3}$.

1.6 Calculate the wave impedances of the electron plasmas in question 1.5 at a frequency of 1 GHz.

1.7 Calculate the electron cyclotron frequency at magnetic field strengths of 0.05, 0.1 and 0.2 T.

1.8 Calculate the propagation constants at a frequency of 500 MHz for positive and negative circularly polarized waves in an electron plasma whose electron density is $10^{16}\,m^{-3}$ in the presence of a magnetic field of 0.01 T.