

Chapter 2

Electromagnetic Waves

The mind of man has perplexed itself with many hard questions. Is space infinite, and in what sense? Is the material world infinite in extent, and are all places within that extent equally full of matter? Do atoms exist or is matter infinitely divisible?

James Clerk Maxwell

Abstract The topic of electromagnetism is extensive and deep. Nevertheless, we have endeavoured to restrict coverage of it to this chapter, largely by focusing only on those aspects which are needed to illuminate later chapters in this text. For example, the Maxwell equations, which are presented in their classical flux and circulation formats in Eqs. (2.1)–(2.4), are expanded into their integral forms in Sect. 2.2.1 and differential forms in Sect. 2.3. It is these differential forms, as we shall see, that are most relevant to the radiation problems encountered repeatedly in ensuing chapters. The process of gathering light from the sun to generate ‘green’ power generally involves collection structures (see Chap. 8) which exhibit smooth surfaces that are large in wavelength terms. The term ‘smooth’ is used to define a surface where any imperfections are dimensionally small relative to the wavelength of the incident electromagnetic waves, while ‘large’ implies a macroscopic dimension which is many hundreds of wavelengths in extent. Under these circumstances, electromagnetic wave scattering reduces to Snell’s laws. In this chapter, the laws are developed fully from the Maxwell equations for a ‘smooth’ interface between two arbitrary non-conducting media. The transverse electromagnetic (TEM) wave equations, which represent interfering waves at such a boundary, are first formulated, and subsequently, the electromagnetic boundary conditions arising from the Maxwell equations are rigorously applied. Complete mathematical representations of the Snell’s laws are the result. These are used to investigate surface polarisation effects and the Brewster angle. In the final section, the Snell’s laws are employed to examine plane wave reflection at perfectly conducting boundaries. This leads to a set of powerful yet ‘simple’ equations defining the wave guiding of electromagnetic waves in closed structures.

2.1 Electromagnetic Spectrum

The study of solar power collection methods is predominantly an exercise in understanding the nature of electromagnetic waves and also in harnessing this widely applicable technology to facilitate the designing, and the optimisation, of optical gathering processes and structures for solar power systems. That light is a form of electromagnetic wave was arguably first established in 1862–1864 by James Clerk Maxwell. The concise set of equations which he developed to explain electromagnetic phenomena (see Sect. 2.2) both predicted the existence of electromagnetic waves and furthermore that these waves would travel with a speed that was very close to the contemporaneously known speed of light. The inference he then made was that visible light and also, by analogy, invisible infrared and ultraviolet rays all represented propagating disturbances (or radiation) occasioned by natural, abrupt changes in electromagnetic fields at some locality in space, such as in the sun. Radio waves, on the other hand, were first detected not from a natural source, but from a wire aerial, into which time-varying currents were deliberately and artificially inserted, from a relatively low-frequency oscillatory circuit. The feat was achieved by the German scientist Heinrich Hertz in 1887.

It is now well established that light (see Fig. 2.1) forms a very small portion of a spectrum of electromagnetic waves which extend from very low-frequency (VLF, MF, VHF at <1 MHz) radio waves, through broadcast waves between 50 and 1,000 MHz, microwaves from 1 to 100 GHz, millimetre waves at about 0.1–1 THz, followed by infrared. The visible spectrum seems narrow when located in the entire electromagnetic spectrum, as presented in Fig. 2.1, but it still encompasses a huge frequency range from 0.43×10^{15} to 0.75×10^{15} Hz (430–750 THz). Beyond the visible section are the ultraviolet, the X-ray and gamma-ray spectra, with a notionally terminal frequency, in an engineering context, for the whole EM spectrum at about 10^{19} Hz which translates to a miniscule wavelength of $0.1 \text{ \AA} = 0.01 \text{ nm}$. Sub-angstrom dimensions are so far outside of normal engineering practice that we need not consider, any further, EM waves at this extremity of the spectrum.

2.2 Electromagnetic Theory and Maxwell's Equations

In traditional electrical engineering science [1, 2], at the macroscopic level where quantum mechanical influences are generally insignificant, all electrical phenomena can be interpreted as being evolved from the forces acting between stationary, or moving, ‘point’ charges (electrons and protons). In fact, four concise equations, commonly referred to as the Maxwell equations, are sufficient to describe all known macroscopic field interactions in electrical science including behaviours at optical frequencies. These equations are in a minimalist mathematical form:

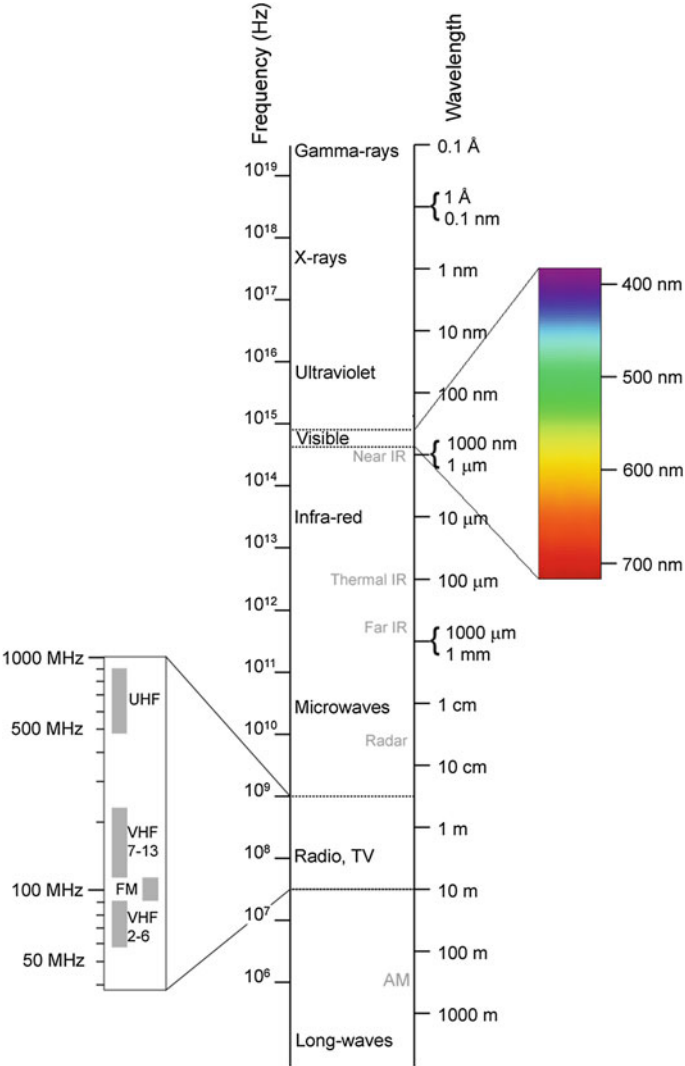


Fig. 2.1 Electromagnetic spectrum (<http://en.wikipedia.org/wiki/File:Electromagnetic-Spectrum.png>)

Flux **D** = charge enclosed (2.1)

Flux **B** = 0 (2.2)

Circ **H** = *I* + rate of change electric flux (2.3)

Circ **E** = −rate of change of magnetic flux (2.4)

Although concisely expressed as they are here, these four equations can, at first sight, still seem rather mystifying, and perhaps a little off-putting, to anyone proposing to study the subject. However, once the symbols and the terminology are established, and the historical development is explained, their obscurity should disappear as their potency is revealed. To this end, the following symbol identifications and further definitions are appropriate.

1. Electric charge (Q), which may be either positive or negative, is conserved in all electrical operations.
2. The electric current through any surface is the rate at which charge passes through the surface, that is, $I = \rho v A$, where ρ is the charge density in coulomb/ m^3 , v is the velocity of moving charge in m/s , while A is the area in m^2 through which charge is passing. The velocity v and the surface normal are presumed to be aligned. The dimensions of I are coulomb/s, which is an amp in the m.k.s. system.
3. The electric current through any *closed* surface is minus the rate of change of the charge enclosed within the surface ($I = -dQ/dt$). This is a general statement of Kirchoff's law, which for circuit engineers mainly appears in the more familiar form $\Sigma I = 0$ at a network junction.

For the purposes of solar engineering, the basic carrier of charge, namely the electron, is considered to be a particle, essentially because the electron wave function in quantum mechanics displays an extremely small wavelength, $\lambda_e = 0.165 \text{ nm}$. This is too short to be detected by engineering instruments. Consequently, the electron's wavelike behaviour is rarely encountered in engineering applications, even in those encompassing optical interactions.

In Eqs. (2.1)–(2.4), the vector quantities $\mathbf{E}$ and $\mathbf{B}$ (the bold type denotes a vector) are the fundamental electric and magnetic field quantities in electromagnetism, while $\mathbf{D}$ and $\mathbf{H}$ are auxiliary fields. Materials embedded within electrical systems are defined electrically by three parameters, namely conductivity σ (mhos $\cdot \text{m}$), permittivity ϵ (Farad/m) and permeability μ (Henry/m). All of these quantities are defined and dimensioned more comprehensively in Ref. [1].

2.2.1 Flux and Circulation

The flux and circulation integrals embedded in Eqs. (2.1)–(2.4) can be defined, rather helpfully, in a relatively non-mathematical form, if averaging (in essence integration) can be considered to be a process which is not unduly remote from 'common sense' (see Refs. [3, 4]). Thus, we have for an arbitrary vector $\mathbf{A}$:

$$\begin{aligned}
 \text{Flux } \mathbf{A} &= \text{average normal component of } \mathbf{A} \text{ over a} \\
 &\quad \text{surface area } d\mathbf{S} \text{ (} A_n \text{ say) multiplied} \\
 &\quad \text{by area } dS \\
 &= A_n \times dS
 \end{aligned}
 \tag{2.5}$$

The introduction of vector algebra into Eq. (2.5) permits a 'shorthand' representation of the process. For an infinitesimally small area (dS), which can be considered (see Fig. 2.2) to be directionally aligned with a unit vector $\mathbf{n}$ normal to its surface, then a simple dot product gives

$$\text{Flux } \mathbf{A} = \mathbf{A} \cdot d\mathbf{S} \tag{2.6}$$

For a surface area S of finite size, we then have [4]

$$\text{Flux } \mathbf{A} = \sum_S \mathbf{A} \cdot d\mathbf{S} = \iint_S \mathbf{A} \cdot d\mathbf{S} \tag{2.7}$$

If the surface of interest is not open, as above, but closed like the surface of a balloon, then Eq. (2.7) takes the form:

$$\text{Flux } \mathbf{A} = \oiint_S \mathbf{A} \cdot d\mathbf{S} \tag{2.8}$$

The mathematical form of circulation can be constructed in a similar manner from the basic definition (Fig. 2.3):

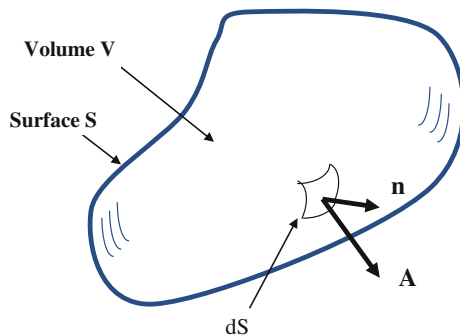


Fig. 2.2 The closed surface S defines the volume V . The direction of the elemental surface $d\mathbf{A}$ is defined by the unit vector $\mathbf{n}$, and vector $\mathbf{A}$ represents in magnitude and direction an arbitrary field passing through it

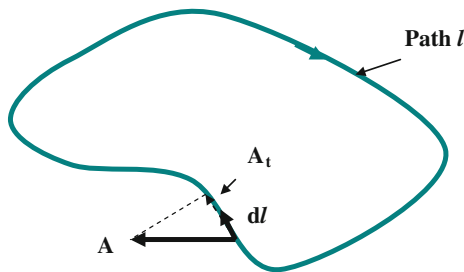


Fig. 2.3 Circulation of $\mathbf{A}$ around the path ℓ is the line integral of the tangential component A_t

$$\begin{aligned} \text{Circ } \mathbf{A} &= \text{average tangential component of } \mathbf{A} \text{ along} \\ &\quad \text{path } d\ell \text{ (} A_t \text{ say) times the length of path } d\ell \\ &= A_t \times d\ell \end{aligned} \quad (2.9)$$

In vector notation, the ‘circulation’ (or perhaps it should be ‘translation’ for an open path) along the elemental path $d\ell$ is given by

$$\text{Circ } \mathbf{A} = \mathbf{A} \cdot d\ell \quad (2.10)$$

For an arbitrary path of length ℓ , circulation is expressed mathematically in the form:

$$\text{Circ } \mathbf{A} = \int_{\ell} \mathbf{A} \cdot d\ell \quad (2.11)$$

For a closed path or loop, which is much more common in electrical calculations, we get

$$\text{Circ } \mathbf{A} = \oint_{\ell} \mathbf{A} \cdot d\ell \quad (2.12)$$

With the above vector definitions in place, we can now express the Maxwell equations in their vector integral form as follows:

$$\oiint_S \mathbf{D} \cdot d\mathbf{S} = Q_{\text{free}} \quad (2.13)$$

$$\oiint_S \mathbf{B} \cdot d\mathbf{S} = 0 \quad (2.14)$$

$$\oint_C \mathbf{H} \cdot d\boldsymbol{\ell} = I_{\text{cond}} + \frac{\partial}{\partial t} \iint_A \mathbf{D} \cdot d\mathbf{A} \quad (2.15)$$

$$\oint_C \mathbf{E} \cdot d\boldsymbol{\ell} = -\frac{\partial}{\partial t} \iint_A \mathbf{B} \cdot d\mathbf{A} \quad (2.16)$$

In Eq. (2.13), Q_{free} denotes the free, unbounded charge within the closed surface S , while in Eq. (2.15), I_{cond} denotes the conducting current, or free charge passing through the open surface A which spans the circuital path C . That is, for positive charge flow,

$$I_{\text{cond}} = \iint_A \rho \mathbf{v} \cdot d\mathbf{A} \quad (2.17)$$

The second term on the right of Eq. (2.15) is Maxwell's displacement current which also threads through the surface A .

Finally, it is important to note that in electromagnetism, a fundamental force equation linking fields and charge also exists as almost a fifth Maxwell equation. It is attributed to Lorentz [5] and is defined in the next section.

2.2.2 Boundary Conditions

At a 'smooth' interface between two different materials (say samples 1 and 2) where smooth implies that surface roughness features are very much less than the free-space wavelength at the frequency of interest, the above four equations reduce to the following boundary conditions:

$$\hat{\mathbf{n}} \cdot \mathbf{D}_1 = \hat{\mathbf{n}} \cdot \mathbf{D}_2 \quad (2.18)$$

$$\hat{\mathbf{n}} \cdot \mathbf{B}_1 = \hat{\mathbf{n}} \cdot \mathbf{B}_2 \quad (2.19)$$

$$\hat{\mathbf{n}} \times \mathbf{E}_1 = \hat{\mathbf{n}} \times \mathbf{E}_2 \quad (2.20)$$

$$\hat{\mathbf{n}} \times \mathbf{B}_1 = \hat{\mathbf{n}} \times \mathbf{B}_2 \quad (2.21)$$

If material 2 is a 'good conductor', the following forms apply:

$$\hat{\mathbf{n}} \cdot \mathbf{D}_1 = \rho_s \quad (2.22)$$

$$\hat{\mathbf{n}} \cdot \mathbf{B}_1 = 0 \quad (2.23)$$

$$\hat{\mathbf{n}} \times \mathbf{E}_1 = 0 \quad (2.24)$$

$$\hat{\mathbf{n}} \times \mathbf{H}_1 = \mathbf{J}_s \quad (2.25)$$

In these equations, $\hat{\mathbf{n}}$ is the unit normal to the surface, ρ_s is the charge density on the surface, and $\mathbf{J}_s$ is the surface current density.

In electromagnetism, the fundamental force equation attributed to Lorentz [3] can be expressed in vectorial form as follows:

$$[\mathbf{F} = Q(\mathbf{E} + \mathbf{u} \times \mathbf{B})] \quad (2.26)$$

In the m.k.s. system, we already know that force is expressed in newtons, Q in coulombs and velocity $\mathbf{u}$ in m/s. In this system, therefore, electric field has the dimension newton/coulomb (N/C), while magnetic flux density $\mathbf{B}$ has the dimension N·s/m·C. Needless to say, we do not use these clumsy forms. In the m.k.s. system, electric field has the basic dimension volt/m, while magnetic flux density gets the dimension tesla (T). The relationship between a volt/m and an N/C and between a tesla and an N·s/m·C can be found in Ref. [1].

2.3 Plane Wave Solution

All materials contain electric charges bound loosely or otherwise within atoms and molecules. If these materials exist in an environment which naturally or artificially causes agitation of the charge, and hence changes in the associated electric and magnetic fields, then electromagnetic waves are unavoidable. These can appear in quite complex trapped, surface, evanescent and radiant embodiments. In these circumstances, the integral forms of Maxwell's equations, developed above, become inappropriate since the finite volumes, surfaces and paths over which integrations have to be performed are no longer identifiable. What is required in this case is a set of equations which represent the field behaviour at a point in space. The conversion from the integral forms to these point forms (differential forms) of Maxwell's equations is developed in most textbooks on the topic (see References) and essentially entails the recruitment of well-known vector-differential theorems such as the divergence theorem and Stokes' law to accomplish the transitions.

Many solar power gathering problems are of the source-free variety, which implies that the source, in this case the sun, is so far distant that the waves of interest here on earth are plane waves. These waves, also termed TEM waves, are described as 'plane' because the radius of curvature of the wave front (see Fig. 1.5) is very large, and thus, the natural rate of curvature of the front can be deemed mathematically insignificant, allowing it to be fully described by means of Cartesian coordinates. In this scenario, the EM problem reduces to a boundary value problem, for which Maxwell's equations, in differential form, become

$$\nabla \cdot \mathbf{D} = 0 \quad (2.27)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.28)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.29)$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \quad (2.30)$$

where $\mathbf{E}$ and $\mathbf{H}$ represent the electric and magnetic field intensities in the region of interest. As before, $\mathbf{D} = \epsilon \mathbf{E}$ is the electric flux density, while $\mathbf{B} = \mu \mathbf{H}$ is the magnetic flux density. The ‘del’ operator (∇) expresses directional derivatives in the three space directions. It is a vector, which in the Cartesian system (for example) has the form:

$$\nabla = \hat{\mathbf{a}}_x \frac{\partial}{\partial x} + \hat{\mathbf{a}}_y \frac{\partial}{\partial y} + \hat{\mathbf{a}}_z \frac{\partial}{\partial z} \quad (2.31)$$

where $\hat{\mathbf{a}}_x$, $\hat{\mathbf{a}}_y$ and $\hat{\mathbf{a}}_z$ are unit vectors directed along x , y , and z , respectively. When the del operator is multiplied by a scalar [$\phi(x, y, z)$ say], the result is a vector which expresses the gradient or slope of the function ϕ in all three space directions, i.e.

$$\nabla \phi = \hat{\mathbf{a}}_x \frac{\partial \phi}{\partial x} + \hat{\mathbf{a}}_y \frac{\partial \phi}{\partial y} + \hat{\mathbf{a}}_z \frac{\partial \phi}{\partial z} \quad (2.32)$$

Cross multiplication of del with a vector produces the operation of ‘curl’, while dot multiplication produces the operation of ‘divergence’ (‘div’). Crudely, curl is circulation at a point, while divergence is flux at a point.

2.3.1 Second-Order Differential Equation

To solve the Maxwell equations for E -field or H -field behaviour in a bounded region, it is first necessary to form an equation either E or H alone. The standard procedure for achieving this conversion is to perform a curl operation on either the curl equation for E or the corresponding equation for H . This gives, for example, using Eq. (2.29)

$$\begin{aligned} \nabla \times \nabla \times \mathbf{E} &= -\frac{\partial}{\partial t} \mu \nabla \times \mathbf{H} \\ &= -\frac{\partial}{\partial t} \left[\frac{\partial}{\partial t} \mu \epsilon \mathbf{E} \right] \\ &= -\mu \epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \end{aligned} \quad (2.33)$$

Hence, on using a convenient vector identity, which states that for any vector $\mathbf{A}$,

$$\nabla \times \nabla \times \mathbf{A} = \nabla \nabla \cdot \mathbf{A} - \nabla^2 \mathbf{A} \quad (2.34)$$

Equation (2.33) can be re-expressed as follows:

$$\nabla \nabla \cdot \mathbf{E} - \nabla^2 \mathbf{E} = -\mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.35)$$

But, from Eq. (2.27), $\nabla \cdot \mathbf{E} = 0$, for a linear, homogeneous medium for which μ and ϵ are constants. Therefore,

$$\nabla^2 \mathbf{E} = \mu\epsilon \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.36)$$

and by analogy:

$$\nabla^2 \mathbf{H} = \mu\epsilon \frac{\partial^2 \mathbf{H}}{\partial t^2} \quad (2.37)$$

Equations (2.36) and (2.37) are wave equations. Equations of this nature, with appropriate variables, appear in most branches of science and engineering, and their solutions have been studied widely. Solutions depend very much on the boundary conditions, namely the conditions imposed on the variables at the periphery or containing surface of the solution region. They can fix the magnitude of the variable (Dirichlet condition) or the rate of change of the variable (Newman condition) or a mixture of both. A unique solution depends on the conditions being neither underspecified or overspecified.

For example, let us consider formulating a solution to Eq. (2.36), and inevitably Eq. (2.37) because of the Maxwell linkages, for a region of free space ($\mu = \mu_0$; $\epsilon = \epsilon_0$) which is large enough to presume that all boundaries are effectively at infinity. In this case, we can choose to represent the region mathematically using Cartesian coordinates, and furthermore, since we anticipate that the solution is a waveform, we can arbitrarily determine that the waves travel in the z -direction. This implies that the rates of change of the E -field in x and y are zero, and using (2.27), it follows that $E_z = 0$. The equation to be solved, therefore, is

$$\frac{\partial^2 \mathbf{E}}{\partial \mathbf{z}^2} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.38)$$

where, in general, $\mathbf{E} = \hat{\mathbf{a}}_x E_x + \hat{\mathbf{a}}_y E_y$. However, if we choose to align the coordinate system so that $\mathbf{E}$ lies along the x -axis (x -polarised solution), then $E_y = 0$ and the wave equation reduces to the scalar form:

$$\frac{\partial^2 E_x}{\partial \mathbf{z}^2} = \frac{1}{c^2} \frac{\partial^2 E_x}{\partial t^2} \quad (2.39)$$

if $c = 1/\sqrt{\mu_0\epsilon_0}$.

2.3.2 General Solution

Equation (2.39) has a wave solution of the general form:

$$E_x = Af(z - ct) + Bf(z + ct) \quad (2.40)$$

This is easily demonstrated by substitution back into the equation. The first term represents a wave travelling in the $+z$ -direction, while the second allows for a reflected wave, if such exists. Given that velocity is the rate of change of z with respect to time, it is evident that c represents velocity (actually phase velocity) of the electromagnetic wave in ‘free space’. For vacuum, it is equal to 3×10^8 m/s. The application of Maxwell’s equations also gives $H_z = 0$ and

$$H_y = \frac{A}{\eta}f(z - ct) + \frac{B}{\eta}f(z + ct) \quad (2.41)$$

Also,

$$\frac{E_x}{H_y} = \pm \sqrt{\frac{\mu_0}{\epsilon_0}} = \pm \eta \quad (2.42)$$

η is termed the free-space wave impedance which for air or vacuum has the value $120\pi \Omega$. The resultant solution is a plane electromagnetic wave, also termed a TEM wave, for which E and H are transverse to the direction of propagation and orthogonal to each other. E and H are also in time phase, as Eq. (2.42) attests (see Fig. 2.4).

Electrical engineers are generally very familiar with the relationship between power (P), voltage (V) and current (I) in the form:

$$P = \frac{1}{2} VI \quad \text{W} \quad (2.43)$$

where V and I are defined in peak, rather than in the more common r.m.s., format. But, voltage is simply integrated electric field E (V/m), and from ampere, current is integrated magnetic field intensity H (A/m), so by analogy, we can suggest that for the plane wave,

$$p = \frac{1}{2} EH = \frac{1}{2} c\epsilon_0 E^2 \quad \text{W/m}^2 \quad (2.44)$$

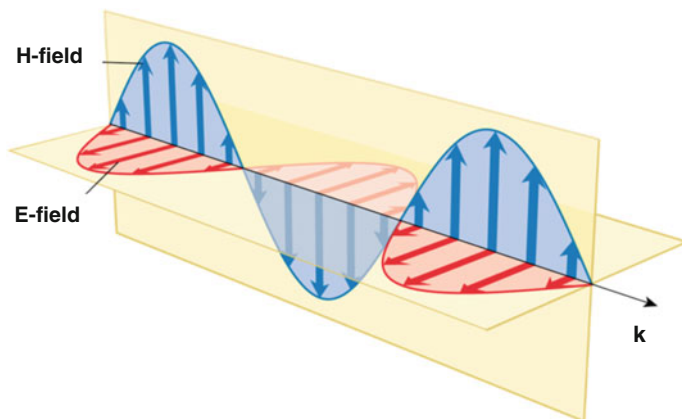


Fig. 2.4 TEM wave field and direction relationships

This means that p is the real power flow density in the TEM wave. In general, complex power flow density in an electromagnetic wave is given by the Poynting vector $\mathbf{S}$, where

$$\mathbf{S} = \frac{1}{2} \mathbf{E} \times \mathbf{H} \quad \text{W/m}^2 \quad (2.45)$$

In electrical engineering, it is much more usual to examine wave solutions at a single frequency (ω rad/s), namely sinusoidal solutions. This actually incurs little loss of generality, since any arbitrary time variation carried on a radio wave can be resolved into a spectrum of single-frequency components. The adoption of a single frequency, or a spectral frequency, in carrying through time-varying computations has the distinct advantage that the time variable can be omitted. The calculations are then progressed in phasor notation. In trigonometric form, Eq. (2.41) becomes

$$E_x = A \exp j(\omega t - \beta z) + B \exp j(\omega t + \beta z) \quad (2.46)$$

where A and B are complex constants. The phasor form is

$$E_x = |A| \exp(-j\beta z + \varphi) + |B| \exp(j\beta z + \theta) \quad (2.47)$$

with φ and θ representing the phases, respectively, of A and B .

2.3.3 Snell's Laws

When a plane electromagnetic wave at the frequency of light, or in fact any radio frequency, is incident upon a smooth interface (by ‘smooth’, it is meant that any surface undulations or protuberances are in size very much less than the wavelength of the impinging waves) between two extended propagating media, part of the wave is reflected back into the incident medium, while part is transmitted or refracted into the second medium, usually with a change of direction.

Analytically, the relationships between the incident and reflected waves can be developed by considering a plane electromagnetic wave, incident at a physically real angle θ_1 to the normal, at the interface between two semi-infinite regions of space, as suggested in Fig. 2.5. Each region is presumed to comprise a linear homogeneous medium with a different index of refraction (n). The index of refraction is defined as follows:

$$n = \frac{c}{v} \quad (2.48)$$

where c is the speed of light in vacuum, or free space, while v is its speed within the specified medium. Also, with reference to Fig. 2.5, the following definitions apply:

$$c = \frac{1}{\sqrt{\mu_0 \epsilon_0}} \quad (2.49)$$

and

$$v_1 = \frac{1}{\sqrt{\mu_0 \epsilon_0 \epsilon_{r1}}} \quad (2.50)$$

$$v_2 = \frac{1}{\sqrt{\mu_0 \epsilon_0 \epsilon_{r2}}} \quad (2.51)$$

Here, ϵ_0 and μ_0 are the free-space permittivity and permeability, respectively, while ϵ_{r1} is the relative permittivity of medium 1 and ϵ_{r2} is the relative permittivity of medium 2. Both media are assumed to be lossless and non-magnetic in which case $\mu_1 = \mu_2 = \mu_0$. The indices of refraction for the two media then become

$$n_1 = \sqrt{\epsilon_{r1}}; \quad n_2 = \sqrt{\epsilon_{r2}} \quad (2.52)$$

Maxwell's equations in the semi-infinite regions remote from the interface are, as we have seen above, fully satisfied by TEM plane waves. It remains then to satisfy the Maxwell boundary conditions at the interface. If this can be done, the resultant solutions represent complete EM solutions for the specified boundary value problem. For an incident TEM wave, as depicted in Fig. 2.5, the E -field vector and the H -field vector must be mutually orthogonal to each other and to the

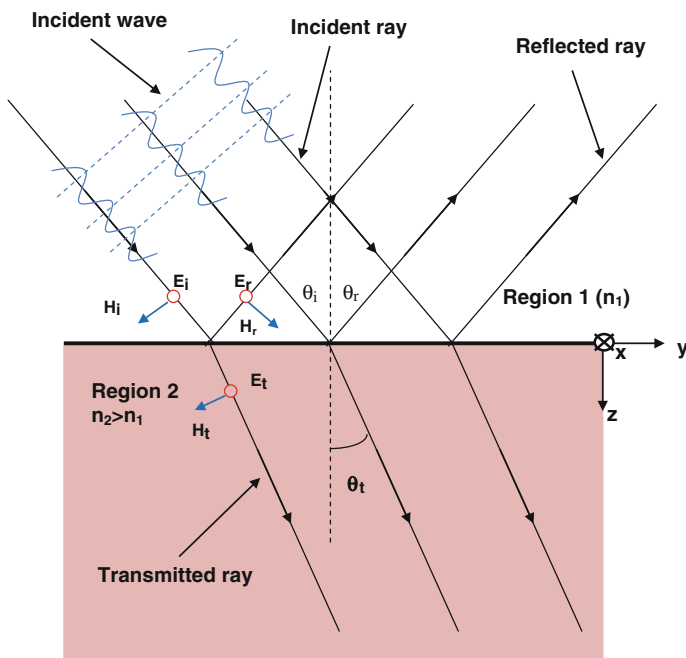


Fig. 2.5 Reflection and refraction at a dielectric interface—perpendicularly polarised case

direction of propagation, usually defined by a unit vector $\hat{\mathbf{k}}$, directed in the direction of the relevant ray.

In this case, we can write

$$\mathbf{H} = \frac{1}{\eta}(\hat{\mathbf{k}} \times \mathbf{E}) \quad (2.53)$$

where η is the wave impedance for the medium containing the wave. Hence, for regions 1 and 2, respectively,

$$\eta_1 = \sqrt{\frac{\mu_0}{\epsilon_{r1}\epsilon_0}} \quad (2.54)$$

$$\eta_2 = \sqrt{\frac{\mu_0}{\epsilon_{r2}\epsilon_0}} \quad (2.55)$$

However, this condition does not fully establish the polarisation direction, which must also be specified. There are two basic choices from which any other polarisation possibilities can be deduced. We can choose the E -field vector of the incident wave to be either normal to the yz -plane, or parallel to it. The yz -plane in Fig. 2.5 is generally termed the plane of incidence for the incoming wave, being the plane that contains

both the direction vector $\hat{\mathbf{k}}$ and the unit normal to the interface ($\hat{\mathbf{n}}$). When the electric field in the incident TEM wave is normal to the plane of incidence, the wave is said to be perpendicularly polarised, and when it is parallel to this plane, it is described as parallel polarised. Note that in relation to the surface of the earth, while parallel polarisation equates to horizontal polarisation, perpendicular polarisation can be termed vertical polarisation only if θ_i approaches 90° . Perpendicular polarisation is often termed transverse electric (TE) propagation, while parallel polarised waves get the complementary description of transverse magnetic (TM) waves.

Now that we know the electromagnetic field forms (TEM waves) remote from the interface between regions 1 and 2 in Fig. 2.5, we can examine the field conditions (boundary conditions) precisely at the interface. For the diffraction set-up depicted in Fig. 2.5 with a perpendicularly polarised TEM wave incident at θ_i , the field directions at a given instant in time can be represented vectorially as shown. Just at the interface, a typical ray of the incident TEM wave is both reflected off the surface and transmitted through it. Also, for a ‘smooth’ surface, ‘common sense’ suggests that it is safe to presume that the reflected and transmitted waves retain the polarisation of the incident wave. Furthermore, there will be a single reflected ray and a single transmitted ray. Actually, this latter assumption is not strictly necessary as we will show presently.

When the TEM wave direction (or ray) lies in paths other than along the coordinate axes, it is usual to define the ray direction by the vector $\mathbf{k}$ which is chosen to be equal in magnitude to the wave coefficient k . That is $\mathbf{k} = \hat{\mathbf{k}}k$. Hence, we can express mathematically the wave component in any other direction ($\mathbf{r}$ say). For the case shown in Fig. 2.6, where the electric field is x -directed, the expression has the form:

$$\mathbf{E}_i = \hat{\mathbf{a}}_x E_i \exp(-j\mathbf{k} \cdot \mathbf{r}) \quad (2.56)$$

Consequently, if $\mathbf{r}$ and $\mathbf{k}$ lie in the yz -plane as suggested in Fig. 2.6, then clearly

$$\mathbf{k} = \hat{\mathbf{a}}_y k_y + \hat{\mathbf{a}}_z k_z \quad (2.57)$$

$$\mathbf{r} = \hat{\mathbf{a}}_y y + \hat{\mathbf{a}}_z z \quad (2.58)$$

Also,

$$k^2 = k_y^2 + k_z^2 \quad (2.59)$$

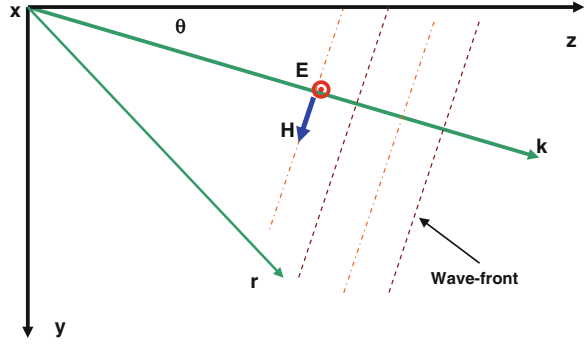
so we can conveniently write

$$k_y = k \sin \theta \quad (2.60)$$

and

$$k_z = k \cos \theta \quad (2.61)$$

Fig. 2.6 Representation of TEM wave with $\mathbf{E}$, $\mathbf{H}$ and $\mathbf{k}$ in mutually orthogonal directions



Hence, employing these relationships, Eq. (2.56) can be expanded into the non-vectorial form:

$$E_{xi} = -E_i \exp(j(\omega t - k_1 z \cos \theta_i - k_1 y \sin \theta_i)) \quad (2.62)$$

where

$$k_1 = \frac{\omega}{v_1} = \frac{\omega}{c} n_1 \quad (2.63)$$

For a TEM wave, the electric and magnetic fields are related through Eq. (2.53). Hence, on combining Eqs. (2.62) and (2.53), and observing the field directions in Fig. 2.5, we obtain for magnetic fields:

$$H_{yi} = -H_i \cos \theta_i \exp[j(\omega t - k_1 z \cos \theta_i - k_1 y \sin \theta_i)] \quad (2.64)$$

$$H_{zi} = H_i \sin \theta_i \exp[j(\omega t - k_1 z \cos \theta_i - k_1 y \sin \theta_i)] \quad (2.65)$$

Also, we note that if these field components represent a TEM wave, then we must have

$$\frac{E_i}{H_i} = \eta_1 = \frac{\eta_0}{n_1} \quad (2.66)$$

Similar constructions lead to the following equations for the reflected and transmitted field components:

$$E_{xr} = -E_r \exp(j(\omega t + k_1 z \cos \theta_i - k_1 y \sin \theta_i)) \quad (2.67)$$

$$H_{yr} = H_r \cos \theta_i \exp[j(\omega t + k_1 z \cos \theta_i - k_1 y \sin \theta_i)] \quad (2.68)$$

$$H_{zr} = H_r \sin \theta_i \exp[j(\omega t + k_1 z \cos \theta_i - k_1 y \sin \theta_i)] \quad (2.69)$$

with

$$\frac{E_r}{H_r} = \eta_1 \quad (2.70)$$

and

$$E_{xt} = -E_t \exp[j(\omega t - k_2 z \cos \theta_t - k_2 y \sin \theta_t)] \quad (2.71)$$

$$H_{yt} = -H_t \cos \theta_t \exp[j(\omega t - k_2 z \cos \theta_t - k_2 y \sin \theta_t)] \quad (2.72)$$

$$H_{zt} = H_t \sin \theta_t \exp[j(\omega t - k_2 z \cos \theta_t - k_2 y \sin \theta_t)] \quad (2.73)$$

where

$$\frac{E_t}{H_t} = \eta_2 \quad (2.74)$$

and

$$k_2 = \frac{\omega}{c} n_2 \quad (2.75)$$

The above field expressions for the incident and reflected waves in region 1 and the transmitted waves in region 2 each separately satisfy Maxwell's equations in these regions. A solution that satisfies Maxwell's equations for the entire volume including the interface is achieved by enforcing the electromagnetic field boundary conditions, given in Eqs. (2.18)–(2.22), at the interface. That is, at $z = 0$, we require that across the divide between regions 1 and 2:

$$E_x \text{ is continuous} \quad (2.76)$$

$$H_y \text{ is continuous} \quad (2.77)$$

$$B_z \text{ is continuous} \quad (2.78)$$

On combining Eq. (2.76) with the field expressions (2.62), (2.67) and (2.71), we obtain with little difficulty:

$$E_{xi} + E_{xr} = E_{xt} \quad (2.79)$$

on the $z = 0$ plane. The implication is that

$$E_i \exp(-jk_1 y \sin \theta_i) + E_r \exp(-jk_1 y \sin \theta_r) = E_t \exp(-jk_2 y \sin \theta_t) \quad (2.80)$$

This equation must remain true over the entire $z = 0$ boundary, from $-\infty \leq y \leq +\infty$. This is only possible if

$$k_1 \sin \theta_i = k_1 \sin \theta_r = k_2 \sin \theta_t \quad (2.81)$$

It is pertinent to note here that if at the commencement of this derivation, we had, without pre-knowledge of refraction rules, chosen to presume that several reflected waves at angles $\theta_{r1}, \theta_{r2}, \theta_{r3} \dots$, and several transmitted waves at angles $\theta_{t1}, \theta_{t2}, \theta_{t3} \dots$, were possible, then the equivalent form of Eq. (2.80) would lead to

$$k_1 \sin \theta_{r1} = k_1 \sin \theta_{r2} = k_1 \sin \theta_{r3} = \dots \dots \dots \quad (2.82)$$

and

$$k_2 \sin \theta_{t1} = k_2 \sin \theta_{t2} = k_2 \sin \theta_{t3} = \dots \dots \dots \quad (2.83)$$

These equations clearly dictate that $\theta_{r1} = \theta_{r2} = \theta_{r3} = \dots$, and $\theta_{t1} = \theta_{t2} = \theta_{t3} = \dots$. In other words, an ‘optically smooth’ surface produces only one reflected wave and one transmitted wave.

Equation (2.81) is the source of Snell’s laws which state that at an optically smooth interface between two lossless media,

$$\theta_r = \theta_i \quad (2.84)$$

$$\frac{\sin \theta_t}{\sin \theta_i} = \frac{k_1}{k_2} = \frac{n_1}{n_2} \quad (2.85)$$

However, these laws govern only the reflection and refraction angles. We also need to have knowledge of the relative magnitudes of the reflected and transmitted waves, and how these are influenced by material properties.

When Eqs. (2.76)–(2.78) are applied to the TEM field components at the boundary, while also applying Snell’s laws, the following relations are generated:

$$E_i + E_r = E_t \quad (2.86)$$

$$(H_i - H_r) \cos \theta_i = H_t \cos \theta_t \quad (2.87)$$

$$(B_i - B_r) \sin \theta_i = B_t \sin \theta_t \quad (2.88)$$

Equation (2.86) can be converted to magnetic field form by employing Eqs. (2.70) and (2.74) leading to

$$\eta_1(H_i + H_r) = \eta_2 H_t \quad (2.89)$$

Consequently, if we choose to define reflection coefficient for this perpendicularly polarised example (TE case) as

$$\rho_{TE} = \frac{H_r}{H_i} \quad (2.90)$$

then making use of Eqs. (2.87) and (2.89), the following useful relationship is deduced:

$$\rho_{\text{TE}} = \frac{\eta_2 \cos \theta_i - \eta_1 \cos \theta_t}{\eta_2 \cos \theta_i + \eta_1 \cos \theta_t} \quad (2.91)$$

This can also be expressed in a slightly more familiar form, which explicitly incorporates the indices of refraction, namely

$$\rho_{\text{TE}} = \frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t} \quad (2.92)$$

Similarly, if we choose to define the transmission coefficient as

$$\tau_{\text{TE}} = \frac{H_t}{H_i} \quad (2.93)$$

then

$$\tau_{\text{TE}} = \frac{2n_2 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_t} \quad (2.94)$$

It is not difficult to demonstrate that

$$\rho_{\text{TE}} = \frac{H_r}{H_i} = \frac{E_r}{E_i} \quad (2.95)$$

and

$$\frac{E_t}{E_i} = \frac{n_1}{n_2} \tau_{\text{TE}} \quad (2.96)$$

An analogous derivation can also be followed through for the parallel polarisation case (TM case). If this is done, we obtain

$$|\rho_{\text{TM}}| = \frac{E_r}{E_i} = \frac{H_r}{H_i} = \left| \frac{n_2 \cos \theta_i - n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t} \right| \quad (2.97)$$

and

$$\tau_{\text{TM}} = \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t} \quad (2.98)$$

The reflection coefficient, as a function of incident angle for both TE and TM cases, is plotted in Fig. 2.7. Clearly, for an interface between lossless dielectrics of differing refractive indices, the reflection behaviours are distinct. While for the TE

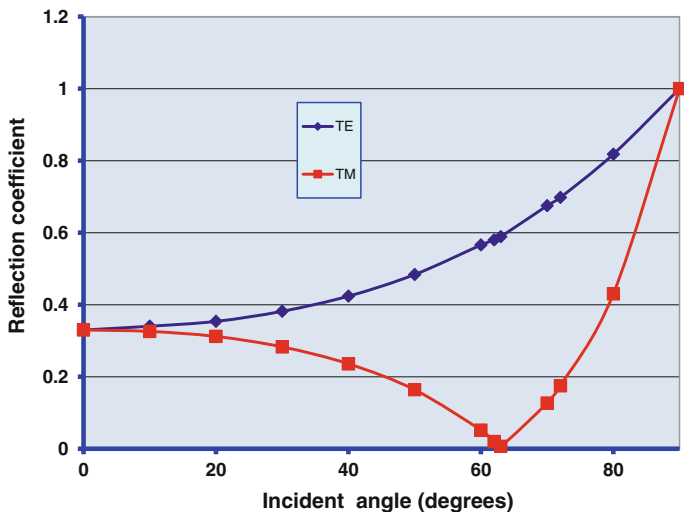


Fig. 2.7 Perpendicular (*diamonds*) and parallel (*squares*) polarised reflection coefficients as a function of incident angle ($n_2 > n_1$)

case, it increases monotonically from a magnitude of 0.33 ($n_1 = 1$ and $n_2 = 2$) at $\theta_i = 0$, to unity at $\theta_i = 90^\circ$, it drops to zero close to 60° in the TM case. At the zero reflection angle, the two surfaces are said to be ‘matched’ for surface-normal wave components. It is termed the Brewster angle, a physical property which underpins the design of light polarisers.

2.3.4 Wave Guiding

Snell’s laws can also be used to explore the processes behind electromagnetic wave trapping or guidance, concepts which are needed in later chapters. While it is well understood that at low frequencies, TEM waves can be guided by a pair of conductors, such as in power lines, in parallel wire telephone lines or in coaxial lines, high-frequency wave trapping in hollow conducting pipes is not so easy to comprehend. Such waveguides are increasingly being used in many of the antenna configurations employed in solar power applications. This method of guidance is very efficient and is especially applicable to high power transmission [2, 7–10]. It relies on the nature of plane wave interference patterns and can, perhaps, best be explained by consideration of Fig. 2.8.

Figure 2.8 depicts (in two dimensions for simplicity) a pair of plane electromagnetic waves (TEM waves) of equal magnitude travelling in different directions A and B. The waves are represented by their wave fronts, with the wave peaks in each case denoted by solid transverse lines (planes in 3D) and wave troughs by dashed lines. The distance between a wave peak and wave trough is, of course, half

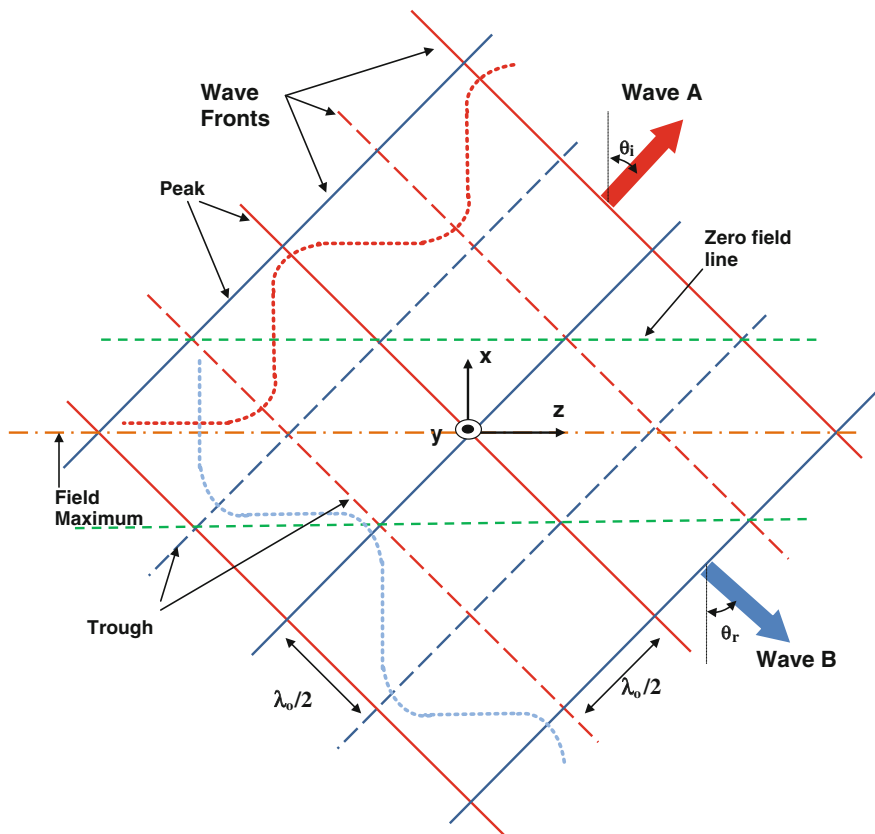


Fig. 2.8 Plane wave interference pattern

of the free-space wavelength ($\lambda_0/2$). The waves are travelling at the velocity of light (c) in the directions of the large arrows. On examination of this wave pattern, it is not too difficult to observe that along the horizontal chain-dotted line (or in three dimensions—the yz -plane), peaks of wave A coincide with those of wave B, and troughs coincide with troughs—and this is independent of the movement of the waves. This line (or plane) represents a stationary (in the x -direction) field maximum ‘independent of time’, while the waves continue to exist.

In contrast, along the green z -directed dashed line, peaks of wave A coincide with troughs of wave B, and vice versa, resulting in a stationary field null at these positions. Consequently, if a perfectly conducting sheet of infinite extent, orientated normal to the x -axis, is located at the stationary null position, the field pattern remains unchanged. For a sheet at the upper dashed line, the red direction arrow (wave A) then represents an incident wave and the blue arrow (wave B) a reflected wave, which, according to Snell’s laws at a perfect mirror, reflects with a magnitude equal to the incident wave and at an angle such that $\theta_r = \theta_i$, as is required to retain the pattern. For perpendicularly polarised plane waves with the E -field confined to

the y -direction, the E -field pattern forms a cosine distribution between the null planes. This pattern can be trapped or guided by introducing a second conducting sheet at the lower null locus in Fig. 2.8. The trapped pattern travels in the z -direction with a phase velocity:

$$v_p = c/\sin \theta \quad (2.99)$$

and a wavelength

$$\lambda_p = \lambda_o/\sin \theta \quad (2.100)$$

where c is the speed of light, λ_o is the TEM wavelength in free space and $\theta = \theta_i = \theta_r$. The magnetic field distribution can easily be deduced by applying trigonometrical rules, and the total E/H pattern is termed a TE guided wave. A dual TM guided wave can be formed by commencing with parallel polarised TEM components.

A TE wave between parallel conducting planes separated by a distance a is illustrated in Fig. 2.9c. The sinusoidal field variations in x are clearly shown. The relationship between plane separation a and wavelength λ_o can again be deduced from trigonometry and yields a

Fig. 2.9 Dominant TE_{10} mode in rectangular waveguide (solid green vectors = E -field; dashed blue = H -field). **a** Side view (TE_{10}). **b** End view (TE_{10}). **c** Top view (TE_{10})

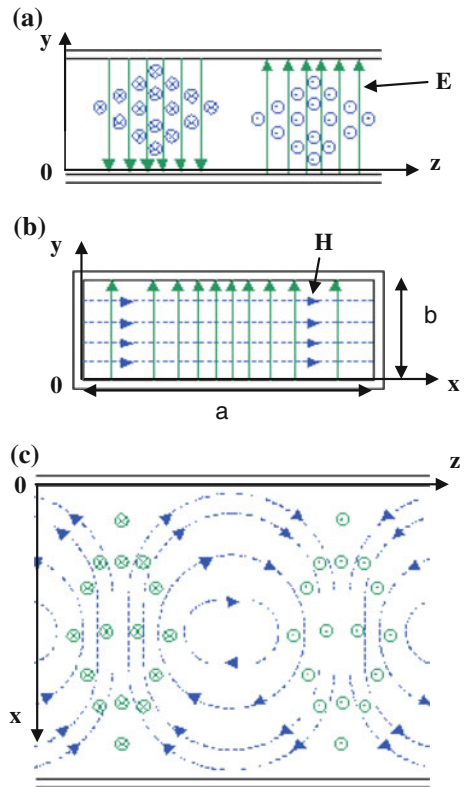
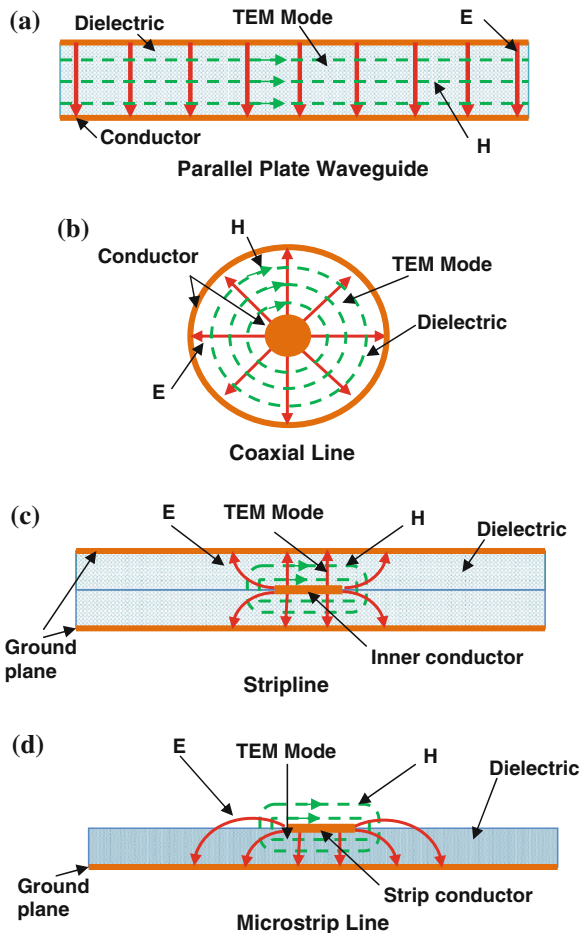


Fig. 2.10 TEM-mode transmission lines



$$\cos \theta = \frac{m\lambda_o}{2a} \quad (2.101)$$

m is the number of half-sinusoids of field pattern between the null planes. In Fig. 2.8, it is not necessary to choose the nearest null planes to create a trapped pattern. Equation (2.101) only has meaning for $m\lambda_o < 2a$, so that for $m = 1$, the case depicted in Fig. 2.9c, the free-space wavelength must be less than $2a$ for propagation to occur. The corollary is that the frequency of the wave $f (=1/\lambda_o)$ must be greater than a certain critical value or cut-off value corresponding to the cut-off wavelength $\lambda_c = 2a$. Furthermore, if $a < \lambda_o < 2a$ the $m = 2, 3, 4$ —solutions all yield the impossible requirement that $\cos \theta > 1$. This means that in the prescribed frequency range, only the $m = 1$ solution is possible. The solution is termed the dominant mode for the parallel plane waveguide of separation a and is defined as

the TE_{10} mode—with one E -field variation in x and zero variation in y . This mode is shown in Figs. 2.9a–c. Perfectly conducting ‘lids’ can be introduced at $y = 0$ and $y = b$ to form a rectangular waveguide, without altering the pattern, because the E -field is normal to these walls. The b -dimension is usually chosen to be approximately half the a -dimension to maximise bandwidth. (For further elucidation, see Sect. 6.5.)

It is relevant to emphasise here that guided electromagnetic waves can also be procured by simply trapping the TEM wave [6] between conductors that lie normal to the electric field vectors, thus satisfying the boundary conditions. The most common alternatives are shown in Fig. 2.10. Parallel plate waveguide (Fig. 2.10a) provides only limited guidance in the direction normal to the plates, but clearly shows how the insertion of smooth conducting planes has negligible effect on the propagation conditions for the TEM mode. Full trapping is provided by coaxial line (Fig. 2.10b) but at the expense of phase velocity reduction and the potential for power loss in the dielectric which is necessary to separate the inner from the outer conductor. Stripline (Fig. 2.10c) is essentially ‘flattened’ coaxial line and has the advantage of ease of fabrication using printed circuit board (PCB) techniques. In coaxial line and in stripline, the dielectric separator usually displays a relative permittivity of between 2 and 3. By increasing this to between 6 and 10 in microstrip (Fig. 2.10d), it becomes possible to dispense with the upper ground plane and create an open structure into which microwave components can relatively easily be inserted.

References

1. Ferrari R (1975) An introduction to electromagnetic fields. Van Nostrand Reinhold Co., Ltd., New York
2. Hammond P (1971) Applied electromagnetism. Pergamon Press Ltd., Oxford
3. Feynman RP, Leighton RB, Sands M (1972) Lectures in physics, vol. II. Addison-Wesley Publishing Co., London
4. Morse PM, Feshbach H (1953) Methods of theoretical physics. McGraw-Hill Book Co., New York
5. Johnk CTA (1988) Engineering electromagnetic fields and waves. Wiley, New York
6. Kraus JD (1984) Electromagnetics. McGraw-Hill Book Co., London
7. Lorrain P, Corson D (1962) Electromagnetic field and waves. W.H. Freeman & Co., San Francisco
8. Baden-Fuller AJ (1993) Engineering electromagnetism. Wiley, New York
9. Bevensee RM (1964) Slow-wave structures. Wiley, New York
10. Stratton JA (2007) Electromagnetic theory. Wiley, New Jersey

Electromagnetic Foundations of Solar Radiation
Collection

A Technology for Sustainability

Sangster, A.J.

2014, XVI, 277 p. 143 illus., Hardcover

ISBN: 978-3-319-08511-1