Fundamentals of Electromagnetic Waves

André Knoesen and Carl Arft

Revised: April 30, 2009

Contents

1 Mathematics for Electromagnetic Waves **1**
1.1 Complex Numbers . . . 3
1.1.1 The “Complex Plane” . . . 3
1.1.2 Complex Conjugate . . . 4
1.1.3 Powers and Roots . . . 4
1.2 Vectors . . . 5
1.2.1 Real Vectors . . . 5
1.2.2 Complex Vectors . . . 8
1.3 Scalar and Vector Fields . . . 8
1.3.1 Gradient, Divergence, and Curl . . . 9
1.3.2 Geometrical Shapes and Vector Analysis . . . 13
1.3.3 Vector Theorems . . . 14
1.3.4 The Spherical Coordinate System . . . 15
1.4 Phasors . . . 17
1.5 Time Averages of Time-harmonic Vectors . . . 20
1.5.1 The Average of a Scalar Wave . . . 20
1.5.2 The Average of the Cross Product of Two Vectors . . . 21
1.6 Important Concepts . . . 21

2 Wave Propagation in Unbounded Media **23**
2.1 The Electromagnetic Spectrum . . . 25
2.2 Fields, Charges, and Currents . . . 25
2.3 Maxwell’s Equations . . . 26
2.4 Wave Propagation in Free Space . . . 27
2.5 Plane Waves in Free Space . . . 32
2.5.1 Vector Phasors . . . 33
2.5.2 Why is a Plane Wave Called a “Plane” Wave? . . . 36
2.5.3 Properties of Plane Waves in Free Space . . . 36
2.6 Wave Propagation in General Media . . . 38
2.6.1 Characteristics of General Media . . . 38
2.6.2 Polarizable and Magnetizable Media . . . 39
2.6.3 Plane Waves in Lossless Media . . . 40
2.6.4 Plane Waves in Lossy Media . . . 42
2.6.5 Example of a Lossy Medium: Wave Propagation in Seawater . . . 46
2.6.6 Plane Waves in a Good Dielectric . . . 47
2.6.7 Plane Waves in a Good Conductor . . . 48
2.6.8 Plane Waves in Dispersive Media . . . 49
2.7 Energy Flow and the Poynting Vector . . . 50
2.8 Polarization . . . 54
2.8.1 Linearly Polarized Waves . . . 56

2.8.2 Circularly Polarized Waves . . . 57
2.8.3 Elliptically Polarized Waves . . . 58
2.9 Important Concepts . . . 59

3 Reflection and Transmission of Waves **63**
3.1 Boundary Conditions . . . 65
3.2 Normal Incidence on a Perfect Conductor . . . 66
3.3 Normal Incidence on a Dielectric . . . 70
3.4 Normal Incidence on a Conducting Medium . . . 73
3.5 Oblique Incidence of Plane Waves at a Single Interface . . . 77
3.5.1 Wave Propagation in a General Direction . . . 77
3.5.2 Plane of Incidence . . . 77
3.5.3 TE and TM Solutions . . . 78
3.5.4 Tangential Phase Matching at an Interface . . . 79
3.5.5 Tangential Phase Matching: "k-Vector Surfaces" and Snell's Law . . . 81
3.5.6 Single Interface: TE-polarized case . . . 83
3.5.7 Single Interface: TM-polarized Case . . . 85
3.5.8 Single Interface: Unified Formulation . . . 87
3.5.9 Total Internal Reflection . . . 93
3.5.10 Brewster's Angle . . . 93
3.5.11 Example: Reflection from an Air/Glass Interface . . . 94
3.6 Power Flow at an Interface . . . 97
3.6.1 TE Polarization . . . 97
3.6.2 TM Polarization . . . 99
3.6.3 Reflectance and Transmittance . . . 99
3.7 Standing Waves . . . 100
3.8 Plane Waves at Multiple Interfaces . . . 101
3.8.1 Relating Amplitudes Across One Interface . . . 102
3.8.2 Relating Amplitudes Within a Layer . . . 103
3.8.3 Relating Amplitudes Between Multiple Layers . . . 105
3.8.4 Example: Three Layer Case . . . 106
3.9 Important Concepts . . . 108

4 Waveguides **111**
4.1 The "Waveguide Coordinate System" . . . 114
4.2 Waveguiding Condition for a Three-Layered Structure . . . 115
4.3 Parallel Plate Waveguides . . . 117
4.3.1 Propagation Constants of the Guided Modes . . . 117
4.3.2 The "Discretization" of k-space . . . 120
4.3.3 Cutoff Conditions for Parallel Plate Waveguides . . . 120
4.3.4 Modal Behavior Below Cutoff . . . 124
4.3.5 Wavelength and Phase Velocity in the Waveguide . . . 125
4.3.6 Electromagnetic Fields: TE Modes . . . 125
4.3.7 Electromagnetic Fields: TM Modes . . . 126
4.3.8 The TEM Mode . . . 128
4.3.9 Guide Impedance in a Parallel Plate Waveguide . . . 129
4.4 Symmetric Dielectric Planar Waveguides . . . 131
4.4.1 Wave Vectors in the Core and Cladding . . . 131
4.4.2 Propagation Constants of the Guided Modes . . . 133
4.4.3 Cutoff Conditions for Guided Modes . . . 135
4.4.4 Electromagnetic Field Analysis . . . 137
4.4.5 TE Modes . . . 137

4.4.6 TM Modes . . . 141
4.5 Rectangular Waveguides . . . 142
4.5.1 Propagation Constants of the Guided Modes . . . 142
4.5.2 Cutoff Conditions . . . 143
4.5.3 Electromagnetic Modes . . . 144
4.5.4 The Fundamental Mode . . . 145
4.5.5 Why does $a = 2b$ in a Typical Waveguide? . . . 147
4.6 Important Concepts . . . 147

5 Group Velocity and Dispersion 149
5.1 Dispersion of Electromagnetic Signals . . . 150
5.1.1 Phase Velocity vs. Group Velocity . . . 150
5.1.2 The Origins of Dispersion . . . 154
5.2 Material Dispersion: Plasma . . . 154
5.3 Waveguide Dispersion: Parallel Plate Waveguide . . . 157
5.4 Important Concepts . . . 158

6 Radiation 161
6.1 Electrostatics Review . . . 164
6.1.1 Scalar Potential . . . 165
6.1.2 The Electric Dipole . . . 167
6.1.3 Vector Potential . . . 168
6.2 Radiation from Time-Varying Charges . . . 169
6.2.1 The Inhomogeneous Wave Equations . . . 169
6.2.2 Solutions to the Inhomogeneous Wave Equations . . . 170
6.3 Radiation from a Hertzian Dipole . . . 171
6.3.1 Near-field Characteristics . . . 173
6.3.2 Far-field Characteristics . . . 174
6.4 Radiator Gain and Radiation Resistance . . . 175
6.5 Radiation from Arrays of Hertzian Dipoles . . . 176
6.6 Important Concepts . . . 184

A Derivations for Chapter 1 187
A.1 The Average of the Cross Product of Two Vectors . . . 188

B Derivations for Chapter 2 189
B.1 Time-Harmonic Form of Maxwell's Equations . . . 190
B.2 General Solution of the Helmholtz Equation . . . 190
B.3 Orthogonality of Wave Vector and Electric Field . . . 191
B.4 Plane Waves in a Good Conductor . . . 192
B.5 Power Flow in Plane Waves . . . 192
B.5.1 In a Lossless Medium . . . 192
B.5.2 In a Lossy Medium . . . 193

C Derivations for Chapter 3 195
C.1 Reflection Coefficient from a Dielectric, Normal Incidence . . . 196
C.2 Surface Current at the Interface of a Perfect Conductor . . . 196
C.3 TM Reflection and Transmission Coefficients . . . 197
C.3.1 Derivation of Coefficients in Terms of Electric Field . . . 197
C.3.2 Alternate Derivation of Reflection/Transmission Coefficients . . . 198
C.3.3 Alternate Derivation of Coefficient Relationship . . . 198
C.4 Derivation of Brewster's Angle . . . 199
C.5 Power Conservation at an Interface (TE Case) . . . 199

C.6 Power Conservation at an Interface (TM Case) . . . 200

D Derivations for Chapter 4 **201**
D.1 Propagation Constant of a Parallel Plate Waveguide . . . 202
D.2 Guide Wavelength in a Parallel Plate Waveguide . . . 202
D.3 Phase Velocity in a Parallel Plate Waveguide . . . 202
D.4 TE Helmholtz Equation in a Waveguide . . . 203
D.5 Alternative TE Field Derivation in a Planar Waveguide . . . 203

E Derivations for Chapter 5 **205**
E.1 Alternate Derivation of Group Velocity in a Plasma . . . 206
E.2 Group Velocity in a Parallel Plate Waveguide . . . 206

F Derivations for Chapter 6 **207**
F.1 Approximation of $1/R_+$ . . . 208
F.2 Inhomogeneous Vector Wave Equation . . . 209
F.3 Inhomogeneous Scalar Wave Equation . . . 209
F.4 Scalar Potential of an Isolated Oscillating Charge . . . 210

G Glossary of Commonly Used Symbols **213**

Chapter 1

Mathematics for Electromagnetic Waves

James Clerk Maxwell (1831-1879)

"The Maxwell electric field theory is contained in his "wonderful" equations, both invention and synthesis. On their publication in 1864, they appeared as revolutionary as the theory of relativity forty years later. The electric field and the magnetic field are propagated through space in the form of waves, according to the new theory. Thus, the effect due to an electric charge does not appear instantaneously in the entire surrounding space at the moment when the charge that produces it appears in one point. Similarly, the effect of a magnetic field caused by the passage of an electric current only makes itself felt gradually by degrees."

—Pierre L. Aryl in *The Electronic Epoch.*

Maxwell's equations completely describe the field of electromagnetics. Partial differential equations, vectors, and complex analysis describe the propagation of time-harmonic, spatially varying electromagnetic field quantities. Maxwell's equations are:

$$\nabla \times \bar{\mathcal{E}}(\bar{r},t) = -\frac{\partial \bar{\mathcal{B}}(\bar{r},t)}{\partial t} \tag{1.1}$$

$$\nabla \times \bar{\mathcal{H}}(\bar{r},t) = \frac{\partial \bar{\mathcal{D}}(\bar{r},t)}{\partial t} + \bar{\mathcal{J}}(\bar{r},t) \tag{1.2}$$

$$\nabla \cdot \bar{\mathcal{D}}(\bar{r},t) = \rho_v(\bar{r},t) \tag{1.3}$$

$$\nabla \cdot \bar{\mathcal{B}}(\bar{r},t) = 0 \tag{1.4}$$

For the purposes of this chapter the following mathematical observations are made:

1. Maxwell's equations form a system of coupled partial differential equations that involve vector and scalar quantities.

2. The time derivatives in these equations describe the time varying behavior of the fields.

3. The equations are linear partial differential equations, and Fourier analysis can be used to analyze the behavior of complicated signal shapes. In Fourier analysis signals are described by a superposition of time-harmonic signals consisting of sines and cosines. The focus is on the behavior of each periodic signal at a known frequency of oscillation. Because the oscillation frequencies are known, we need to determine how the phase of each signal is being changed by the system of governing equations. Similar to circuit theory, in electromagnetics we use phasors to simplify the manipulation of time-harmonic signals. The phasors are, in general, complex numbers that describe an amplitude and phase shift, but unlike phasors in circuit theory, the phasors are vectors.

4. Analogous to using a superposition of phasors to describe more complicated time-varying signals, in the next chapter we will introduce the concept of a plane wave. A superposition of plane waves can be used to describe more complex *spatially* varying waves. The use of time-harmonic signals in combination with plane waves simplifies the complicated set of Maxwell's coupled partial differential equations into simple ordinary equations with straightforward solutions.

5. The electromagnetic fields are functions of position. The dot products and cross products are spatially dependent derivatives, and it is these derivatives in Maxwell's equations that describe the behavior of the electromagnetic fields in space. At each point in space, the field is assigned either a magnitude alone, or a magnitude and direction. Such fields are referred to as *scalar fields,* $\Phi(\bar{r})$, and *vector fields,* $\bar{A}(\bar{r})$, respectively.

A thorough understanding of Maxwell's equations, then, requires knowledge of complex numbers, phasors, vectors, curl and divergence operations, and time-harmonic signals both in scalar and vector form. A brief review of these fundamental mathematical concepts is given next. You are encouraged to obtain a physical as well as a mathematical interpretation of the concepts.

1.1 Complex Numbers

1.1.1 The "Complex Plane"

A complex number, z, consists of a real and imaginary part, which can be visualized in the "complex plane" as shown in Figure 1.1. The complex number can be thought of as a vector starting at the origin of the complex plane and ending at the coordinate (a,b), where a is the real part and b is the imaginary part of the complex number. The magnitude of this vector, r, is then given by $r = \sqrt{a^2 + b^2}$, and the angle of this vector with respect of the positive real axis, θ, is given by $\theta = \tan^{-1}(b/a)$. The complex number z can therefore be represented in one of two forms. The rectangular form is given by $z = a + jb$, and the polar form is given by $z = re^{j\theta}$, which makes use of Euler's formula:

$$e^{j\theta} = \cos\theta + j\sin\theta \tag{1.5}$$

where $j = \sqrt{-1}$. Equation 1.5 is an important result which we will use repeatedly when transforming between the time- and phasor-domains.

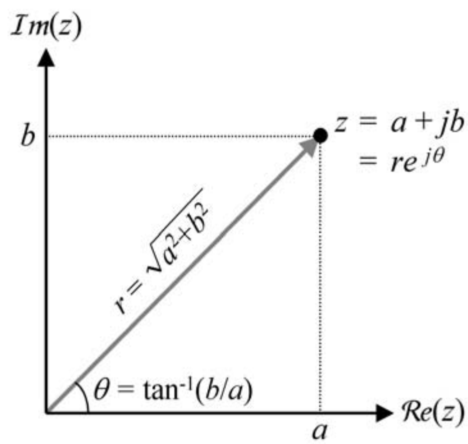


Figure 1.1: Plotting a complex number in the complex plane

1.1.2 Complex Conjugate

The complex conjugate of z is defined in rectangular form as

$$z^* = (a + jb)^* = a - jb \tag{1.6}$$

and in polar form as

$$z^* = \left(re^{j\theta}\right)^* = re^{-j\theta} \tag{1.7}$$

The complex conjugate is a convenient way to extract the real and imaginary part of a complex number z:

$$\mathcal{R}e\{z\} = \frac{z + z^*}{2} = \frac{1}{2}\{a + jb + a - jb\} = a \tag{1.8}$$

$$\mathcal{I}m\{z\} = -j\left(\frac{z - z^*}{2}\right) = -j\frac{1}{2}\{a + jb - (a - jb)\} = b \tag{1.9}$$

1.1.3 Powers and Roots

The n^{th} power of a complex number is best found using the polar form:

$$z^n = r^n e^{jn\theta} = r^n\{\cos(n\theta) + j\sin(n\theta)\} \tag{1.10}$$

Likewise, the n^{th} root of a complex number is given by

$$z^{1/n} = r^{1/n}e^{j\theta/n} = r^{1/n}\left\{\cos\left(\frac{\theta + 2k\pi}{n}\right) + j\sin\left(\frac{\theta + 2k\pi}{n}\right)\right\} \tag{1.11}$$

where $k = 0, 1, 2, ...n-1$. From this it follows that there are n different values for $z^{1/n}$. For example, the square root of a complex number is $z^{1/2} = \pm\sqrt{re^{j\theta}} = \pm\sqrt{r}e^{j\theta/2}$.

$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

EXAMPLE 1.1:

Given the following complex numbers: $A = 1 + j2$ and $B = -2 - j$, calculate the numerical solutions to the following operations: AB, A/B, and $\sqrt{A}$.

Solution:

Carrying out the multiplication in rectangular form:

$$AB = (1 + j2)(-2 - j) = -2 - j - j4 + 2 = -j5$$

or, in polar form:

$$AB = (1 + j2)(-2 - j) = \left(\sqrt{5}e^{j1.107}\right)\left(\sqrt{5}e^{-j2.678}\right) = 5e^{-j1.571} = -j5$$

Calculating the quotient in rectangular form:

$$A/B = \frac{1 + j2}{-2 - j}\left(\frac{-2 + j}{-2 + j}\right) = \frac{-4 - j3}{5} = -0.8 - j0.6$$

or, in polar form:

$$A/B = \frac{\sqrt{5}e^{j1.107}}{\sqrt{5}e^{-j2.677}} = e^{j3.784} = -0.8 - j0.6$$

Finally, calculating the square root in polar form:

$$\sqrt{A} = \left(\sqrt{5}e^{j1.107}\right)^{1/2} = \pm\left(5^{1/4}\right)\left(e^{j1.107/2}\right) = \pm 1.495e^{j0.554} = \pm(1.272 + j0.786)$$

$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

1.2 Vectors

1.2.1 Real Vectors

In this section we first consider only real vectors, i.e. vectors that consist only of real numbers. In a three dimensional, cartesian coordinate system, a real vector is specified by $\bar{V} = V_x\hat{x} + V_y\hat{y} + V_z\hat{z}$ where $\hat{x}$, $\hat{y}$, and $\hat{z}$ are unit vectors along the x-, y-, and z-axes, respectively, and V_x, V_y, and V_z are the magnitudes of the vector $\bar{V}$ in the x-, y-, and z-directions. The length, or norm, of a real vector $\bar{V}$ is given by

$$|\bar{V}| = \sqrt{V_x^2 + V_y^2 + V_z^2} \tag{1.12}$$

The unit vectors $\hat{x}$, $\hat{y}$, and $\hat{z}$ have a length equal to unity, i.e. $|\hat{x}| = |\hat{y}| = |\hat{z}| = 1$.

The addition or subtraction of two vectors $\bar{U} = U_x\hat{x} + U_y\hat{y} + U_z\hat{z}$ and $\bar{V} = V_x\hat{x} + V_y\hat{y} + V_z\hat{z}$ is simply the addition or subtraction of the individual components:

$$\bar{U} \pm \bar{V} = (U_x \pm V_x)\hat{x} + (U_y \pm V_y)\hat{y} + (U_z \pm V_z)\hat{z} \tag{1.13}$$

There are two multiplicative operations for vectors, the dot product and the cross product. The dot product results in a scalar. The cross product results in a vector. The dot product of two real vectors is defined as

$$\bar{U} \cdot \bar{V} = U_x V_x + U_y V_y + U_z V_z = |\bar{U}||\bar{V}| \cos\theta \tag{1.14}$$

where θ is the angle between the two vectors. Notice that we can rewrite the expression for the length of a real vector (Equation 1.12) using the dot product:

$$|\bar{V}| = \sqrt{\bar{V} \cdot \bar{V}} \tag{1.15}$$

Carrying out the dot product, it is straightforward to show that Equations 1.12 and 1.15 are equivalent.

The cross product of two real vectors is defined as

$$\bar{U} \times \bar{V} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ U_x & U_y & U_z \\ V_x & V_y & V_z \end{vmatrix}$$

$$= \hat{x} \begin{vmatrix} U_y & U_z \\ V_y & V_z \end{vmatrix} - \hat{y} \begin{vmatrix} U_x & U_z \\ V_x & V_z \end{vmatrix} + \hat{z} \begin{vmatrix} U_x & U_y \\ V_x & V_y \end{vmatrix} = \hat{n}|\bar{U}||\bar{V}| \sin\theta \tag{1.16}$$

where $\hat{n}$ is a unit vector that is perpendicular to both $\bar{U}$ and $\bar{V}$, and the direction is determined by the right hand rule. The right-hand rule states that the orientation of the vectors' cross product is determined by placing $\bar{U}$ and $\bar{V}$ tail-to-tail, flattening the right hand, extending it in the direction of $\bar{U}$, and then curling the fingers in the direction of the vector $\bar{V}$. The thumb then points in the direction of $\bar{U} \times \bar{V}$, as shown in Figure 1.2.

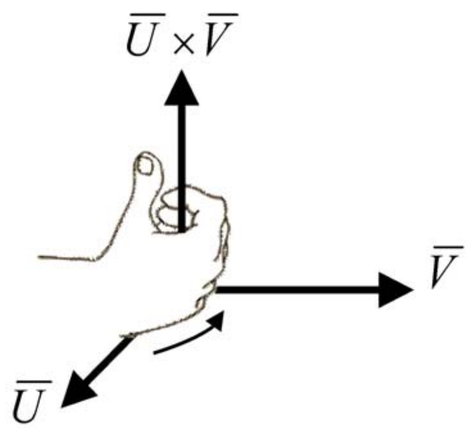


Figure 1.2: Illustration of the right hand rule for the cross product of two vectors. To find the direction of $\bar{U} \times \bar{V}$, start with the fingers of the right hand pointing in the direction of $\bar{U}$ and curl them in the direction of $\bar{V}$. Your thumb will point in the direction of $\bar{U} \times \bar{V}$.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 1.2:

What is the spatial relationship between the vectors $\bar{k}$ and $\bar{E}$, if $\bar{k} \cdot \bar{E} = 0$? (Assume that both $\bar{k}$ and $\bar{E}$ have nonzero length.)

Solution:

From Equation 1.14 we see that if both $\bar{k}$ and $\bar{E}$ have nonzero length, the only way that $\bar{k} \cdot \bar{E} = 0$ is if $\cos\theta = 0$. This occurs when the angle between the vectors is $\theta = 90°$. Therefore, $\bar{k}$ and $\bar{E}$ are orthogonal (i.e. perpendicular) to each other in space.

$\lll\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 1.3:

What is the spatial relationship between the vectors $\bar{k}$ and $\bar{A}$, if $\bar{k} \times \bar{A} = 0$? (Assume that both $\bar{k}$ and $\bar{A}$ have nonzero length.)

Solution:

From Equation 1.16 we see that if both $\bar{k}$ and $\bar{A}$ have nonzero length, then we must have either $\theta = 0°$ or $\theta = 180°$. Therefore, $\bar{k}$ and $\bar{A}$ are either parallel or anti-parallel in space.

$\lll\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 1.4:

Generate a unit vector that is perpendicular to both of the vectors $\bar{A} = 2\hat{x} + 3\hat{y} + \hat{z}$ and $\bar{B} = \hat{x} + \hat{y} + 2\hat{z}$.

Solution:

The cross product generates a vector that is perpendicular to both vectors:

$$\bar{A} \times \bar{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ 2 & 3 & 1 \\ 1 & 1 & 2 \end{vmatrix} = 5\hat{x} - 3\hat{y} - \hat{z}$$

Then, normalize the vector:

$$\hat{n} = \frac{\bar{A} \times \bar{B}}{|\bar{A} \times \bar{B}|} = \frac{1}{\sqrt{35}} \left(5\hat{x} - 3\hat{y} - \hat{z}\right)$$

$\lll\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 1.5:

What is the angle between the vectors $\bar{A} = \hat{x} + 2\hat{y} - 3\hat{z}$ and $\bar{B} = -2\hat{x} + 3\hat{y} - \hat{z}$?

Solution:

We can calculate the angle from the dot product. Solving Equation 1.14 for θ:

$$\theta = \cos^{-1}\left(\frac{\bar{A} \cdot \bar{B}}{|\bar{A}||\bar{B}|}\right) = \cos^{-1}\left(\frac{7}{\sqrt{14}\sqrt{14}}\right) = 60°$$

$\lll\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg\ggg$

1.2.2 Complex Vectors

A complex vector is a vector with one or more components being complex. As we will see, in electromagnetics we often use complex vectors. In many ways, a complex vector is treated similarly to a real vector. For example, vector addition and subtraction are performed in the same manner using element-by-element addition or subtraction (Equation 1.13). The dot and cross products are also done in a similar manner (Equations 1.14 and 1.16) except they can no longer be defined in terms of an angle, θ, since this angle no longer makes physical sense when dealing with vectors in complex space.

An important difference occurs when calculating the length of a complex vector. For example, consider finding the length of the vector $\bar{D} = \hat{x} + j\hat{y}$ using our previous definition (Equations 1.12 and 1.15). We would find $|\bar{D}| = 1^2 + j^2 = 1 - 1 = 0$ which clearly makes no sense. Instead, we use the complex conjugate and define the length of a complex vector as

$$|\bar{V}| = \sqrt{\bar{V} \cdot \bar{V}^*} \tag{1.17}$$

Using this definition, we find the length of vector $\bar{D}$ to be $|\bar{D}| = \sqrt{(\hat{x} + j\hat{y}) \cdot (\hat{x} - j\hat{y})} = \sqrt{1+1} = \sqrt{2}$. Notice that, in the case that $\bar{V}$ is a real vector, Equations 1.17 and 1.15 are equivalent.

1.3 Scalar and Vector Fields

A spatially varying field is described either as a *scalar field,* $\Phi(\bar{r})$, or a *vector field,* $\bar{A}(\bar{r}) = A_x\hat{x} + A_y\hat{y} + A_z\hat{z}$, where the position vector, $\bar{r}$, is given by

$$\bar{r} = x\hat{x} + y\hat{y} + z\hat{z} \tag{1.18}$$

in Cartesian coordinates. This vector describes the current position in space. For example, the point $(x, y, z) = (1, 2, 3)$ is described by the position vector $\bar{r} = \hat{x} + 2\hat{y} + 3\hat{z}$. The difference between a scalar field and a vector field is the following. A scalar field is described by a *magnitude only* at each point in space. A vector field, on the other hand, is described by a *magnitude and a direction* at each point in space, and therefore a vector is required to describe the field at each spatial location $\bar{r}$.

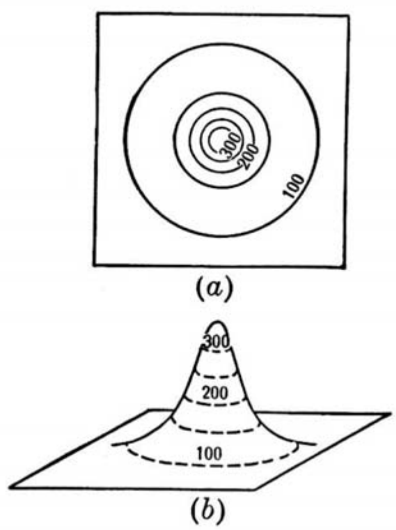


Figure 1.3: (a) Contour map of a mountain. (b) The corresponding mountain in three-dimensions [1].

To understand the difference between scalar and vector fields, consider the following example. Figure 1.3(a) shows a contour map of the mountain shown in Figure 1.3(b). The elevation increases toward the peak, which is at the center of the map. Elevation, as plotted on a contour map, is a *scalar quantity.* Each point on the map corresponds to an elevation of a certain number of feet, and when all points of equal elevation are connected by contour lines, the profile of the earth's surface is completely defined. Elevation is a scalar field.

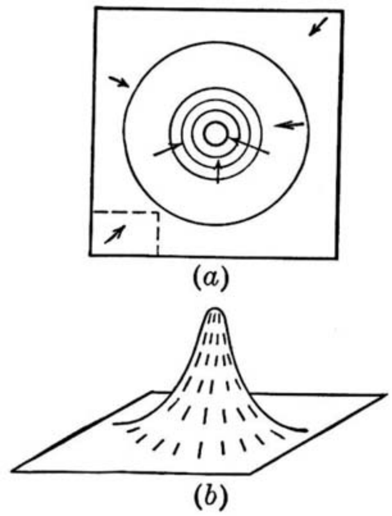


Figure 1.4: (a) Map of a mountain showing the force required at each location to keep a marble in place. (b) The force vectors on the mountain [1].

A marble placed somewhere on the mountain shown in Figure 1.3, gravity forces will make it roll down the hill, in the direction of steepest decent. If we wanted to hold the marble in place, a force would have to be applied in the proper direction on the marble. Depending upon the location of the marble, a different force in a different direction would be required. Figure 1.4(a) shows a map of the mountain on which arrows are drawn to indicate the amount and direction of the force required to hold the marble at various places. The force as plotted on the map is a *vector quantity.* The force to counter act gravity, is a vector field.

1.3.1 Gradient, Divergence, and Curl

The gradient operator ∇ (called "del" or sometimes "nabla"), in Cartesian coordinates,

$$\nabla = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \right) \tag{1.19}$$

is used to describe spatial variations of fields. We use the *gradient* to describe spatial variations in scalar fields and the *divergence* and *curl* to describe variations in vector fields.

Gradient

Gradient is the *magnitude and direction of the greatest rate of change of a scalar field.* The gradient operation is simply the gradient operator multiplied by the scalar field $\Phi(\overline{r})$:

$$\nabla\Phi(\overline{r}) = \frac{\partial\Phi(\overline{r})}{\partial x}\hat{x} + \frac{\partial\Phi(\overline{r})}{\partial y}\hat{y} + \frac{\partial\Phi(\overline{r})}{\partial z}\hat{z} \tag{1.20}$$

The physical interpretation of the gradient is that the *direction* of the gradient is the orientation in which the directional derivative has the largest value, and the *magnitude* of the gradient is the value of that directional derivative. Notice that the gradient of a scalar field is a vector field.

Continuing our previous example of a marble placed on the slope of a mountain, it is apparent that there is a relation between the scalar field of elevation and the vector field of force on a marble. It is a simple and familiar one: the force is dependent upon the steepness of the slope, or, in other words, upon the rate of change of elevation with respect to distance. This rate of change is a derivative, similar in nature to the ordinary derivative of differential calculus. It is complicated, however, by the necessity of finding the direction of steepest decent in order to determine the direction in which the marble will tend to roll. The steepest slope at a given point is the *gradient* of the elevation at that point. It is a vector quantity at each point, and therefore constitutes a vector field. (Adapted from [1]).

Divergence

Divergence is the *rate of change of the field strength in the direction of a vector field* [1]. The divergence of a vector field, $\overline{A}(\overline{r})$, is defined as the dot product of the gradient operator and the vector field:

$$\begin{aligned}\nabla \cdot \overline{A}(\overline{r}) &= \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}\right) \cdot (A_x\hat{x} + A_y\hat{y} + A_z\hat{z}) \\ &= \frac{\partial}{\partial x}A_x + \frac{\partial}{\partial y}A_y + \frac{\partial}{\partial z}A_z \qquad (1.21)\end{aligned}$$

The physical significance of the divergence of a vector field is that it describes the rate at which the "field density" exits a given region of space. Notice that the divergence of a vector field is a scalar field (i.e. a scalar quantity at each point in space).

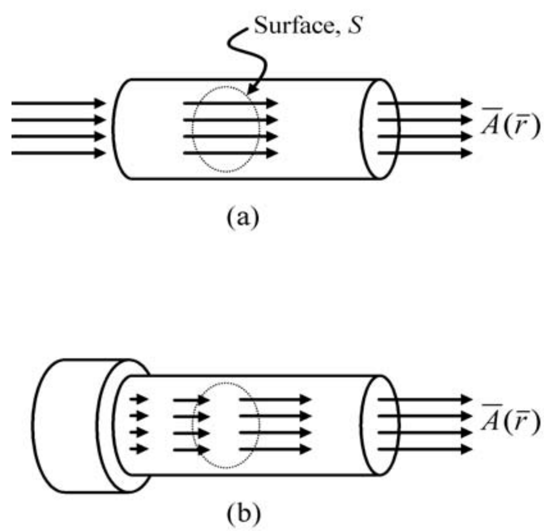


Figure 1.5: Illustration of the divergence of a fields. (a) A pipe through which water is flowing. (b) A tube of compressed air, capped on one end [1].

The divergence of a vector field is nicely illustrated in the flow of fluids. Figure 1.5(a) shows a pipe through which water is flowing. The dashed line within the pipe represents an imaginary surface, S, through which water passes. The surface is completely closed, and water will flow in through one side and out through the other. The water may flow in any irregular fashion, but since water is incompressible the amounts of water entering and exiting the closed surface must be equal. This is the same as saying that the water must flow in such a way that, if its velocity is represented by the vector field $\overline{A}(\overline{r})$, the divergence of $\overline{A}(\overline{r})$ must be zero everywhere. It is from this concept that the name "divergence" arises. Water cannot diverge from any point for it would leave a vacuum; it cannot converge to any point for it is incompressible.

Now consider the flow of air, which is compressible. Figure 1.5(b) represents a tube of compressed air, capped on one end. A similar cap has just been removed from the other end causing air to rush out. Consider the closed surface within the tube that is represented by the dashed line. Because the air is expanding, more air is passing out through the right side of the indicated surface than is entering the surface through the left side. Consequently there is a divergence of air. There is divergence at every point where air is expanding and therefore, if velocity of air is represented by the vector field $\overline{A}(\overline{r})$, the divergence of $\overline{A}(\overline{r})$ is not zero everywhere. (Adapted from [1]).

Curl

Curl is the *rate of change of the field strength in a direction at right angles to the field*, and describes the amount of "rotation" of the vector field [1]. The curl of a vector field, $\overline{A}(\overline{r})$, is defined as the cross product between the gradient operator and the vector field. In Cartesian coordinates, the curl in determinant form is

$$\nabla \times \overline{A}\left(\overline{r}\right) = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

$$= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}\right)\hat{x} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x}\right)\hat{y} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}\right)\hat{z} \tag{1.22}$$

Notice that the curl of a vector field results in another, different vector field.

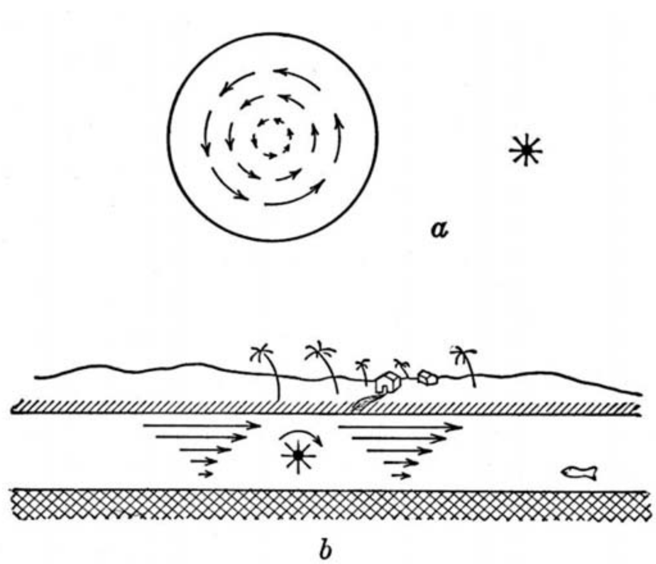


Figure 1.6: Illustration of the curl of a vector field. (a) A tub containing water that has been stirred with a paddle. (b) Water flowing in a canal [1].

To understand the curl of a vector field, consider a tub of water as shown in Figure 1.6(a). The water in it has been stirred with a paddle and the vectors represent the velocity of the water, $\overline{v}$. A small paddle-wheel

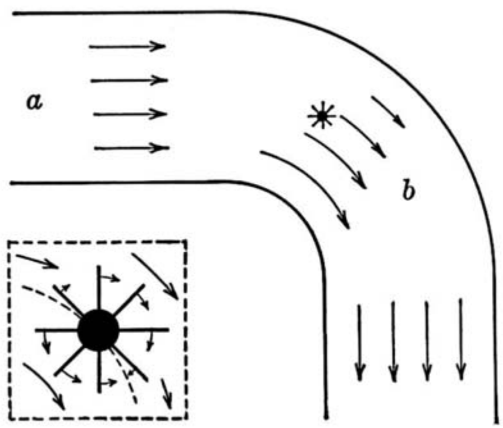


Figure 1.7: Water flowing around a corner in a canal [1].

is shown beside the tub; if this paddle-wheel is mounted on frictionless bearings and dipped into the center of the tub, it will be turned in a counterclockwise direction. At any point the paddle-wheel is placed in the tub, it will be turned by the water, for, even if it is not in the center of the tub, the water will be going more rapidly past one side of the wheel than past the other. The turning of the paddle-wheel is an indication that the water is moving in the tub in such a way that the vector field of velocity has a rate of change, referred to as "curl."

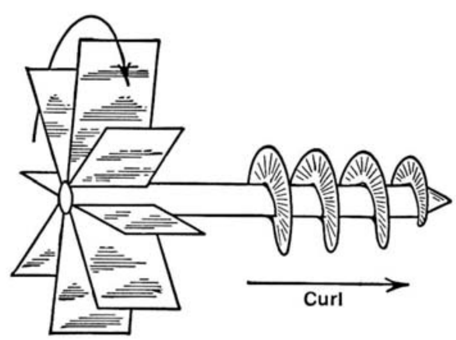


Figure 1.8: The sense of the curl vector is determined by the direction of rotation of the paddle-wheel: if the paddle-wheel turns a right-hand screw, it will screw itself in the direction of the curl vector [1].

The name "curl" indicates an association with motion in curved lines. This is not necessary, however, for straight-line motion may also have curl. If water flows in a canal, as in Figure 1.6(b), in such a way that it flows more rapidly near the surface than it does along the bottom of the canal, every particle of water may move in a straight line, but nevertheless there is a curl, as will be recognized when this vector field with a small paddle-wheel. As indicated in the figure, the paddle-wheel will be turned in a clockwise direction, for the stream is more rapid on the wheel's upper blades than on its lower ones. Figure 1.7 shows a map of another canal, in which the water flows without curl. In the straight part of the canal, the velocity of flow is uniform, and it is obvious that the paddle-wheel at position *a* will not turn. At *b*, in a bend of the canal, it is possible for water to turn the corner without curl, provided it flows faster along the inner margin of the channel in just the right proportion. An enlarged view of the paddle-wheel at *b* is shown (it must be understood that the exploring paddle-wheel is in fact so small that it does not interfere with the flow of water), and little arrows indicate the reaction of the water on each of its blades. Because of the curvature of the lines of flow, more than half of the blades are driven clockwise. But the velocity of water is greater on the inner side, and, although fewer blades are driven counterclockwise, they are each acted on more forcefully. It is readily conceivable that curvature and variation of velocity might be so related that the wheel would have no total tendency to turn. Curved motion is therefore possible without curl. This kind of flow is, as a matter of fact, characteristic of a truly frictionless liquid. It is the purpose of "streamlining" to provide a surface past which air or water will flow with a minimum of curl, for motion with curl develops eddies that waste energy.

The divergence of a vector field is a scalar quantity. There is divergence from a point or to a point (positive or negative), but no idea of direction is involved. The curl of a vector field, on the contrary, is a vector. If curl is visualized as an eddy, it is evident that the eddy must be about some axis; a vertical axis, horizontal axis, or perhaps at some angle. The direction of such an axis is, by definition, the direction of the vector that represents curl. Referring to the hypothetical paddle-wheel, when it is in the position in which it turns most rapidly its axis is in the direction of the curl vector. Each component of the vector of curl may be found by placing the paddle-wheel axis parallel to the appropriate axis of coordinates. The sense of the curl vector is determined by the direction of rotation of the paddle-wheel: if the paddle-wheel turns a right-hand screw, it will screw itself in the direction of the curl vector, as in Figure 1.8. (Adapted from [1]).

1.3.2 Geometrical Shapes and Vector Analysis

In electromagnetics, planes or spheres are used to describe the propagation of waves in space. For this reason, it is important to recognize the vector equations that describe a plane and a sphere. Recall that the general equation for a plane in space is $a_1x + a_2y + a_3z = d$, which can be rewritten in vector form as

$$\bar{A} \cdot \bar{r} = d \tag{1.23}$$

where $\bar{A} = a_1\hat{x} + a_2\hat{y} + a_3\hat{z}$ is a vector that is normal (perpendicular) to the plane, and $\bar{r}$ is the position vector. The geometric interpretation of Equation 1.23 is shown in Figure 1.9.

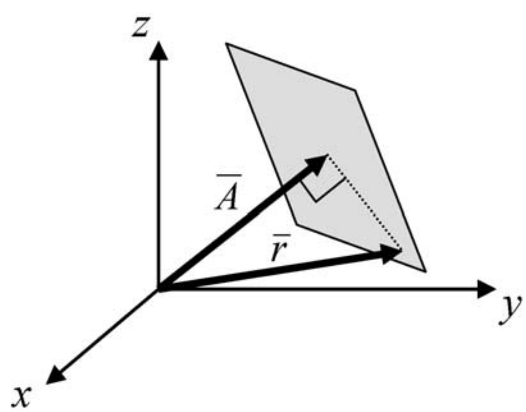


Figure 1.9: If the position vector $\bar{r} = x\hat{x} + y\hat{y} + z\hat{z}$ is constrained to satisfy $\bar{A} \cdot \bar{r} = d$, then $\bar{r}$ will trace out a plane in space.

A sphere of radius R is given by the equation $x^2 + y^2 + z^2 = R^2$. This is equivalent, then, to constraining the position vector $\bar{r}$ such that

$$\bar{r} \cdot \bar{r} = R^2 \tag{1.24}$$

This vector equation will trace out a sphere of radius R. The geometric interpretation of Equation 1.24 is shown in Figure 1.10.

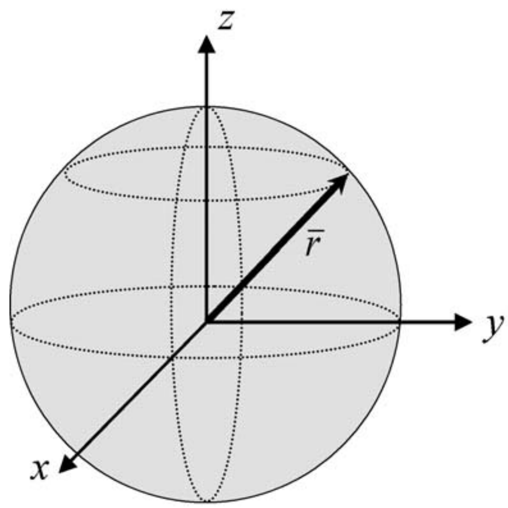


Figure 1.10: If the vector $\bar{r} = x\hat{x} + y\hat{y} + z\hat{z}$ is constrained such $\bar{r} \cdot \bar{r} = R^2$ where R is a constant, the vector $\bar{r}$ will trace out a sphere with radius R.

1.3.3 Vector Theorems

We now present a few vector theorems that we will use in our study of electromagnetics.

Theorem 1: If a vector field is found to have no curl, that vector field is the gradient of some scalar field.

That is,

$$\text{If} \quad \nabla \times \bar{A} = 0, \quad \text{then} \quad \bar{A} = -\nabla \Phi \tag{1.25}$$

Therefore, if a vector field $\bar{A}$ has no curl, there exists a scalar field, Φ, whose gradient is equal to the vector field $\bar{A}$.

One of Maxwell's equations (Equation 1.1) tells us that if the electromagnetic fields do not depend on time, then the curl of the electric field is zero Therefore from this theorem we know that some scalar field can be found, which we will call Φ, such that the negative of the gradient of Φ equals $\bar{E}$. The process of finding Φ may be so difficult that it may not be worth the effort, but we still know that Φ exists. Such a scalar field is called a *scalar potential.*

Theorem 2: If a vector field is found to have no divergence, that vector field is the curl of some other vector field.

That is,

$$\text{If} \quad \nabla \cdot \bar{B} = 0, \quad \text{then} \quad \bar{B} = \nabla \times \bar{A} \tag{1.26}$$

Therefore, if a vector field $\bar{B}$ has no divergence, there exists another vector field $\bar{A}$ whose curl is equal to the vector field $\bar{B}$.

One of Maxwell's equations (Equation 1.4) tells us that the divergence of the magnetic flux density is zero. From this theorem, we know that some other vector field can be found, which we will call $\bar{A}$, such that the curl of $\bar{A}$ is equal to the magnetic flux density. The process of finding $\bar{A}$ may be so difficult that it may not be worth the effort, but we still know that $\bar{A}$ exists. Such a vector field is called a *vector potential.*

Theorem 3: Gauss' Theorem

$$\int_V \left(\nabla \cdot \bar{A}\right) dV = \oint_S \bar{A} \cdot d\bar{S} = \oint_S \left(\bar{A} \cdot \hat{n}\right) dS \tag{1.27}$$

where $\bar{A}$ is a vector field. Gauss' theorem, also called Green's theorem, relates the integral of the divergence within any volume to the integral of the vector field strength over the surface enclosing that volume, where $\hat{n}$ is a unit vector normal to the surface.

Theorem 4: Stokes' Theorem

$$\oint_C \bar{A} \cdot d\bar{l} = \int_S \left(\nabla \times \bar{A}\right) \cdot d\bar{S} \tag{1.28}$$

where $\bar{A}$ is a vector field. Stokes' theorem relates the line integral of a vector along a closed path to the surface integral of the curl of the vector, where the surface is the enclosed by the closed path.

1.3.4 The Spherical Coordinate System

For most of the electromagnetic problems discussed in this course, it makes sense to use the Cartesian coordinate system. One situation where this is not the case, however, is when discussing radiation which is emitted radially outward from the origin. In this case, we will make use of the spherical coordinate system, as shown in Figure 1.11.

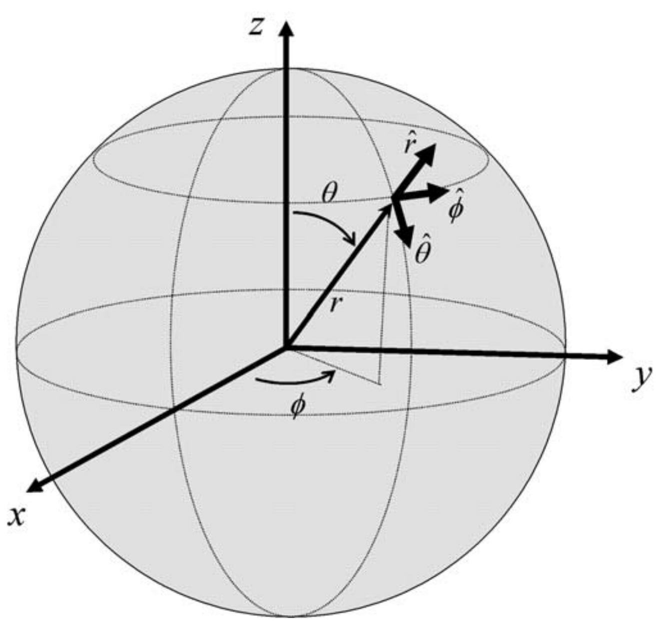


Figure 1.11: Illustration of the spherical coordinate system with unit vectors.

In the spherical coordinate system, the position vector is given by

$$\bar{r} = r\hat{r} + \theta\hat{\theta} + \phi\hat{\phi} \tag{1.29}$$

The unit vectors may be transformed from Cartesian to spherical coordinates as follows:

$$\hat{r} = \hat{x}\sin\theta\cos\phi + \hat{y}\sin\theta\sin\phi + \hat{z}\cos\theta \tag{1.30}$$

$$\hat{\theta} = \hat{x}\cos\theta\cos\phi + \hat{y}\cos\theta\sin\phi - \hat{z}\sin\theta \tag{1.31}$$

$$\hat{\phi} = -\hat{x}\sin\phi + \hat{y}\cos\phi \tag{1.32}$$

and from spherical to Cartesian:

$$\hat{x} = \hat{r}\sin\theta\cos\phi + \hat{\theta}\cos\theta\cos\phi - \hat{\phi}\sin\phi \tag{1.33}$$

$$\hat{y} = \hat{r}\sin\theta\sin\phi + \hat{\theta}\cos\theta\sin\phi + \hat{\phi}\cos\phi \tag{1.34}$$

$$\hat{z} = \hat{r}\cos\theta - \hat{\theta}\sin\theta \tag{1.35}$$

The dot product and cross product of two vectors in spherical coordinates are similar to the dot and cross products in Cartesian coordinates. Given two vectors $\bar{U} = U_r\hat{r} + U_\theta\hat{\theta} + U_\phi\hat{\phi}$ and $\bar{V} = V_r\hat{r} + V_\theta\hat{\theta} + V_\phi\hat{\phi}$ the dot product is

$$\bar{U}\cdot\bar{V} = U_rV_r + U_\theta V_\theta + U_\phi V_\phi \tag{1.36}$$

and the cross product is

$$\bar{U}\times\bar{V} = \begin{vmatrix} \hat{r} & \hat{\theta} & \hat{\phi} \\ U_r & U_\theta & U_\phi \\ V_r & V_\theta & V_\phi \end{vmatrix} \tag{1.37}$$

However, the gradient, curl, and divergence are quite different in spherical coordinates. The gradient operation on a scalar field $\Phi(\bar{r})$ in spherical coordinates is

$$\nabla\Phi(\bar{r}) = \frac{\partial\Phi(\bar{r})}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial\Phi(\bar{r})}{\partial\theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial\Phi(\bar{r})}{\partial\phi}\hat{\phi} \tag{1.38}$$

The divergence of a vector field $\bar{A}(\bar{r}) = A_r\hat{r} + A_\theta\hat{\theta} + A_\phi\hat{\phi}$ in spherical coordinates is

$$\nabla\cdot\bar{A}(\bar{r}) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2A_r\right) + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\left(A_\theta\sin\theta\right) + \frac{1}{r\sin\theta}\frac{\partial A_\phi}{\partial\phi} \tag{1.39}$$

Finally, the curl of a vector field in spherical coordinates in determinant form is

$$\nabla\times\bar{A}\left(\bar{r}\right) = \begin{vmatrix} \frac{1}{r^2\sin\theta}\hat{r} & \frac{1}{r\sin\theta}\hat{\theta} & \frac{1}{r}\hat{\phi} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial\phi} \\ A_r & rA_\theta & r\sin\theta A_\phi \end{vmatrix}$$

$$= \frac{1}{r\sin\theta}\left\{\frac{\partial}{\partial\theta}\left(A_\phi\sin\theta\right) - \frac{\partial A_\theta}{\partial\phi}\right\}\hat{r} - \frac{1}{r\sin\theta}\left\{\frac{\partial}{\partial r}\left(A_\phi r\sin\theta\right) - \frac{\partial A_r}{\partial\phi}\right\}\hat{\theta}$$
$$+ \frac{1}{r}\left\{\frac{\partial}{\partial r}\left(rA_\theta\right) - \frac{\partial A_r}{\partial\theta}\right\}\hat{\phi} \tag{1.40}$$

1.4 Phasors

Time harmonic signals are observed in many situations. In circuit theory we study the response of voltages and currents in a circuit when it is excited with a time-harmonic signal. At any point in the circuit, a single sinusoidal wave describes the current or voltage. The current and voltage are scalar, not vector, quantities. For this reason, we use time-harmonic *scalar* phasors in circuit theory, in which the phasor amplitude is a single complex number and is not a function of position. This is in contrast to electromagnetics, in which the physical quantities are time-harmonic *vector* fields. In electromagnetics, then, we need a *phasor that is also a vector.*

Why do we use phasors? It greatly simplifies the algebra of time-harmonic functions. Consider the time-harmonic signal $\mathcal{V}(t) = V_o \cos(\omega t + \phi)$ where V_o the amplitude of the cosine wave, ω is the angular frequency, and ϕ is the phase offset. This signal is shown in Figure 1.12. The angular frequency ω (radians/sec) of the wave is related to the frequency f (cycles/sec, sec^{-1}, or Hertz) by $\omega = 2\pi f$. In addition, the period of the wave, T, which is defined as the time between adjacent peaks or troughs, is $T = 1/f = 2\pi/\omega$. To derive the phasor representation of this wave, note that using Euler's formula (Equation 1.5) we have

$$\mathcal{V}(t) = V_o \cos(\omega t + \phi) = \mathcal{R}e\left\{V_o e^{j(\omega t+\phi)}\right\}$$

$$= \mathcal{R}e\left\{V_o e^{j\phi} e^{j\omega t}\right\} = \mathcal{R}e\left\{V e^{j\omega t}\right\} \tag{1.41}$$

where $V = V_o e^{j\phi}$. We refer to V as the "phasor amplitude" of $\mathcal{V}(t)$. The phasor amplitude V is a *complex number* that *does not depend on time.* The relationship between the actual wave and the phasor form is

$$\mathcal{V}(t) = V_o \cos(\omega t + \phi) \quad \Longrightarrow \quad V = V_o e^{j\phi} \tag{1.42}$$

To convert from the phasor representation back to the actual wave, we simply multiply the phasor by $e^{j\omega t}$ and then take the real part of the result, as shown in Equation 1.41.

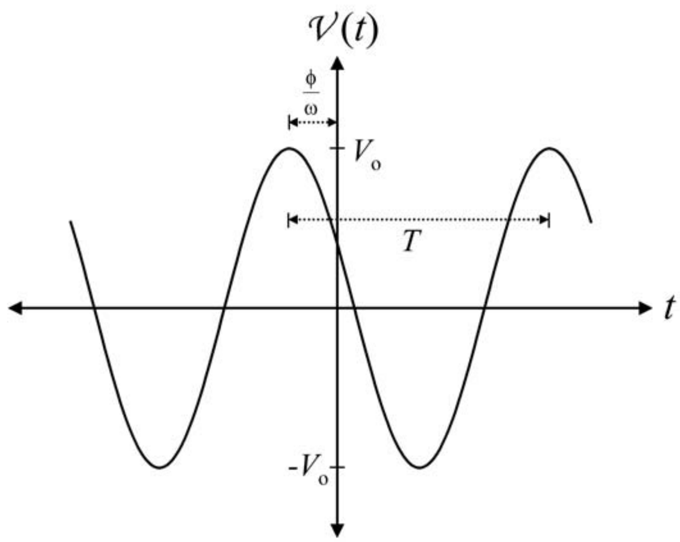


Figure 1.12: Plot of $V_o \cos(\omega t + \phi)$

The behavior of phasors can be thought of in the following way. The phasor V establishes the amplitude and initial phase offset of the wave. When plotted in the complex plane, V would correspond to a vector of length V_o and angle ϕ from the positive real axis. When V is multiplied by $e^{j\omega t}$, it causes the vector to

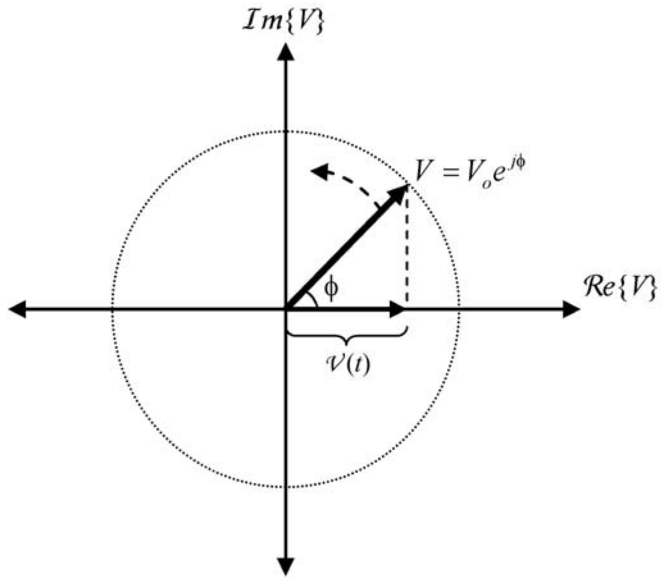


Figure 1.13: Phasor representation in the complex plane. As time increases the phasor rotates at an angular frequency ω. The actual wave is determined by $\mathcal{V}(t) = \mathcal{R}e\{Ve^{j\omega t}\}$, which is the projection of the rotating phasor onto the real axis.

"rotate" at an angular frequency ω. The cosine form of the wave is then determined by the projection of this rotating vector onto the real axis, as shown in Figure 1.13.

As stated previously, using the phasor representation greatly simplifies the manipulation of time-harmonic signals. For example, consider taking the time-derivative of a general cosine signal:

$$\frac{d}{dt}V_o\cos(\omega t + \phi) = -\omega V_o \sin(\omega t + \phi) \tag{1.43}$$

In phasor form, this simplifies to a simple multiplication:

$$\frac{d}{dt}\mathcal{V}(t) \quad \Longrightarrow \quad j\omega V \tag{1.44}$$

This can be shown by explicitly carrying out the derivative using the phasor forms:

$$\begin{aligned}\frac{d}{dt}\mathcal{V}(t) &= \frac{d}{dt}\mathcal{R}e\left\{Ve^{j\omega t}\right\} = \mathcal{R}e\left\{j\omega V e^{j\omega t}\right\} = \mathcal{R}e\left\{j\omega V_o e^{j\phi} e^{j\omega t}\right\} \\ &= \mathcal{R}e\left\{j\omega\, V_o[\cos(\omega t + \phi) + j\sin(\omega t + \phi)]\right\} = -\omega V_o \sin(\omega t + \phi)\end{aligned} \tag{1.45}$$

Electromagnetic waves can propagate in three dimensional space, and are not always constrained in one dimension such as along a wire or circuit. Therefore, we encounter time-harmonic quantities whose amplitude and phase varies in the x-, y-, and z-directions, and must be described by a time-harmonic vector:

$$\begin{aligned}\bar{\mathcal{V}}(t) &= V_{\text{xo}}\cos(\omega t + \phi_x)\hat{x} + V_{\text{yo}}\cos(\omega t + \phi_y)\hat{y} + V_{\text{zo}}\cos(\omega t + \phi_z)\hat{z} \\ &= \mathcal{R}e\left\{\left[V_{\text{xo}}e^{j\phi_x}\hat{x} + V_{\text{yo}}e^{j\phi_y}\hat{y} + V_{\text{zo}}e^{j\phi_z}\hat{z}\right]e^{j\omega t}\right\} = \mathcal{R}e\left\{\bar{V}e^{j\omega t}\right\}\end{aligned} \tag{1.46}$$

For this reason, in electromagnetics we use a *vector phasor* $\bar{V}$, i.e. a phasor that has x-, y-, and z-components. Because electromagnetic fields are vector fields that vary throughout space, the vector phasors are *also* a function of space.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 1.6:
Convert the time-harmonic function $\mathcal{V}(t) = 100\sin(120\pi t + \pi/3)$ to phasor form.

Solution:
The easiest way to convert from the time domain to the phasor domain is to first convert the sine to a cosine:

$$\mathcal{V}(t) = 100\cos(120\pi t + \pi/3 - \pi/2) = 100\cos(120\pi t - \pi/6)$$

Because the wave is now in cosine form, we simply pick out the amplitude and phase:

$$V = 100e^{-j\pi/6}$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 1.7:
What are the harmonics in the signal $\mathcal{V}(t) = \cos(\omega t + \pi/3)\sin(\omega t)$? Can this signal be represented by a single phasor at a frequency ω?

Solution:
The product can be calculated by changing the cosine and sine into exponential form:

$$\mathcal{V}(t) = \left[\frac{e^{j(\omega t+\pi/3)} + e^{-j(\omega t+\pi/3)}}{2}\right]\left(\frac{e^{j\omega t} - e^{-j\omega t}}{j2}\right)$$

$$= \frac{-j}{4}\left\{e^{j(2\omega t+\pi/3)} - e^{-j(2\omega t+\pi/3)} - e^{j\pi/3} + e^{-j\pi/3}\right\}$$

Transforming back into sines and cosines using Euler's formula:

$$\mathcal{V}(t) = \frac{-j}{4}[j2\sin(2\omega t + \pi/3) - j2\sin(\pi/3)] = \frac{1}{2}\sin(2\omega t + \pi/3) - \frac{\sqrt{3}}{4}$$

The signal consists of a DC term, and a harmonic at 2ω. Therefore, it cannot be represented by a phasor at ω.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 1.8:
For the vector phasor $\bar{C} = j\hat{x} + (1+j)\hat{y} + 4\hat{z}$, find the time-harmonic signal $\bar{\mathcal{C}}(t)$.

Solution:
To convert to the time domain, multiply by $e^{j\omega t}$ and take the real part:

$$\bar{\mathcal{C}}(t) = \mathcal{R}e\left\{\bar{C}e^{j\omega t}\right\} = \mathcal{R}e\left\{[j\hat{x} + (1+j)\,\hat{y} + 4\hat{z}]\,e^{j\omega t}\right\}$$

Let us look at each component separately. The x-component is:

$$\mathcal{R}e\left\{je^{j\omega t}\right\} = \mathcal{R}e\left\{j[\cos(\omega t) + j\sin(\omega t)]\right\} = -\sin(\omega t)$$

The y-component is:

$$\mathcal{R}e\left\{(1+j)\,e^{j\omega t}\right\} = \mathcal{R}e\left\{\left(\sqrt{2}e^{j\pi/4}\right)e^{j\omega t}\right\}$$

$$= \mathcal{R}e\left\{\sqrt{2}[\cos(\omega t+\pi/4) + j\sin(\omega t+\pi/4)]\right\} = \sqrt{2}\cos(\omega t+\pi/4)$$

The z-component is:

$$\mathcal{R}e\left\{4e^{j\omega t}\right\} = \mathcal{R}e\left\{4[\cos(\omega t) + j\sin(\omega t)]\right\} = 4\cos(\omega t)$$

The time-harmonic signal is then:

$$\bar{\mathcal{C}}(t) = -\sin(\omega t)\hat{x} + \sqrt{2}\cos(\omega t+\pi/4)\hat{y} + 4\cos(\omega t)\hat{z}$$

$\lll\lll\lll\ggg\ggg\ggg$

1.5 Time Averages of Time-harmonic Vectors

In electromagnetics, the time average of a wave is an important quantity because it determines the power carried by the wave. In this section, we present a mathematical review of the time average of scalar and vector waves.

1.5.1 The Average of a Scalar Wave

The time average of any periodic function, $\mathcal{F}(t)$, is found by integrating over one period and dividing by the period:

$$\langle\mathcal{F}(t)\rangle = \frac{1}{T}\int_0^T \mathcal{F}(t)dt \tag{1.47}$$

Calculating the time average of a simple cosine wave $\mathcal{V}(t) = V_o\cos(\omega t)$, we find that it is zero:

$$\langle\mathcal{V}(t)\rangle = \frac{1}{T}\int_0^T V_o\cos(\omega t)\,dt = 0 \tag{1.48}$$

This is because the wave is symmetric about the time axis, and therefore the positive and negative portions "cancel out" when averaged, as can be seen in Figure 1.12. Assuming that $\mathcal{V}(t)$ represents the voltage across a particular circuit component, it makes physical sense that the average voltage could be zero, since the polarity of the voltage is constantly reversing. But, this does not mean that the *average power* dissipated in the component is zero. This is because the power is proportional to the square of the voltage, which has a DC offset, as shown in Figure 1.14:

$$\mathcal{V}(t)^2 = [V_o\cos(\omega t)]^2 = \frac{V_o^2}{2}[\cos(2\omega t)+1] \tag{1.49}$$

The result is a time average that is nonzero:

$$\left\langle\mathcal{V}(t)^2\right\rangle = \frac{1}{T}\int_0^T \mathcal{V}(t)^2 dt = \frac{V_o^2}{2T}\int_0^T [\cos(2\omega t)+1]\,dt = \frac{V_o^2}{2} \tag{1.50}$$

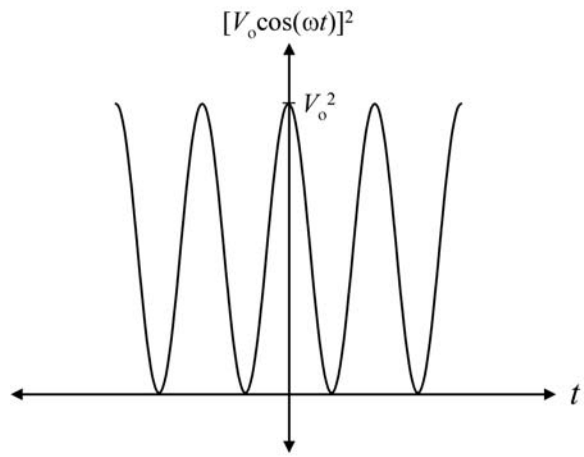


Figure 1.14: The time-harmonic wave $\mathcal{V}(t)^2 = [V_o \cos(\omega t)]^2$. Notice that there is a nonzero average value.

1.5.2 The Average of the Cross Product of Two Vectors

Because we are dealing with vector fields, it becomes more complicated to calculate the time average power of an electromagnetic wave. As we will see, the power in a time-harmonic electromagnetic wave is related to the time average of the *cross product* of two vector phasors. If the two vector fields, $\bar{\mathcal{A}}(t)$ and $\bar{\mathcal{B}}(t)$, have the phasor representations $\bar{A}$ and $\bar{B}$, a very simple way to calculate this time average without having to actually perform the integration is (see Appendix A):

$$\left\langle \bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t) \right\rangle = \frac{1}{T} \int_0^T \left[\bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t) \right] dt = \frac{1}{2} \mathcal{R}e \left\{ \bar{A} \times \bar{B}^* \right\} \tag{1.51}$$

This is an extremely important conclusion. An important quantity (the time average) looked like a formadible mathematical operation, yet if the phaseor are used it is reduced to a simple multiplicative operation.

1.6 Important Concepts

Upon the completion of this chapter, you should be able to:

- Perform addition and subtraction, multiplication, and division of complex numbers.
- Find the square root of a complex number.
- Add and subtract vectors.
- Calculate the dot product and cross product of two vectors, and understand their physical significance.
- Perform the gradient, divergence, and curl operations, and understand their physical significance.
- Recognize the vector description of a plane and a sphere.
- Use and understand the vector theorems.
- Understand why phasors are used in electromagnetics.
- Transform a wave into its phasor form, and return from the phasor form back into the time domain.
- Understand the concept of a vector phasor.
- Quickly calculate the time-average of the cross product of two time harmonic vectors using phasors.

Chapter 2

Wave Propagation in Unbounded Media

"Maxwell made a most important remarkable proposal as follows: It is known by experiment that conduction current produces a magnetic field; total current is for mathematical purposes best expressed as the sum of conduction and displacement current; is it not, then, likely that displacement current also produces a magnetic field? ... This hypothesis led to a conclusion of fundamental importance, for Maxwell showed that, if it were true, energy would be transmitted as electromagnetic waves."

—Hugh H. Skilling in *Fundamentals of Electric Waves*

In this chapter, we first analyze the propagation of waves in free space. Using this as a foundation, we then extend our discussion to understanding wave propagation in more complicated media. In addition, a secondary goal of this chapter is to master the mathematical tools required for manipulating vector fields that vary harmonically in both space and time.

2.1 The Electromagnetic Spectrum

Waves are all around us. Ocean waves travel on the surface of the water. Earthquakes cause waves that travel on the earth's surface. What we perceive as sound is acoustic waves that travel through the air. Ocean waves, earthquake waves, and sound waves all require a medium through which to travel. Sound can travel through air because air is made of molecules, which carry the sound waves by "bumping" into each other, like dominoes knocking each other over. Sound can travel through anything made of molecules; however, there is no sound in outer space because there are no molecules there to transmit the sound waves.

Electromagnetic waves are unlike sound waves in that they do not need a medium through which to travel. Electromagnetic waves can travel through air and solid materials - but they can also travel through empty space. An experiment that beautifully demonstrates the differences between electromagnetic waves and sound waves is the ringing of a bell inside a transparent jar. When air is removed from the jar, the volume of the ringing bell decreases, even though you can still clearly see the bell. This is also why astronauts can see each other but need to use radios to communicatea, can feel heat from the sun, which travels as infrared waves.

The electromagnetic spectrum is illustrated in Figure 2.1. Radio waves, television waves, microwaves, infrared radiation and visible light are all electromagnetic waves. They only differ from each other in wavelength. The wavelength through the electromagnetic spectrum vary from very long radio waves, which have wavelengths measured in hundreds or thousands of meters, to very short gamma rays, which have nanometer or even *picometer* wavelengths (smaller than an atom!).

The wavelength of an electromagnetic wave is inversely proportional to its *frequency*, which is measured in oscillations per second, or Hertz (Hz). For an electromagnetic wave that is travelling through a vacuum, the frequency, f (in Hz), and wavelength, λ (in meters), are related by

$$c = \lambda f \tag{2.1}$$

where the constant $c \approx 2.998 \times 10^8$ m/s is the speed of an electromagnetic wave in a vacuum. This leads to an important observation that *all electromagnetic waves in vacuum travel at the same speed regardless of their wavelength*! Consider the implications if this were not the case.

2.2 Fields, Charges, and Currents

A time-varying electromagnetic field consists of an electric field coupled to a magnetic field. As we will show later, the electric and magnetic fields are perpendicular to each other, and both are perpendicular to the direction of wave propagation. The fields are vector fields (fields with magnitude *and* direction at each point in space) that, in general, vary in space and time. An electromagnetic wave is described by five vector fields:

1. The *electric field intensity,* $\bar{\mathcal{E}}(\bar{r}, t)$, with units of volts per meter (V/m),

2. The *magnetic field intensity,* $\bar{\mathcal{H}}(\bar{r}, t)$, with units of amperes per meter (A/m),

3. The *electric current density,* $\bar{\mathcal{J}}(\bar{r}, t)$, with units of amperes per square meter (A/m^2),

4. The *electric flux density,* $\bar{\mathcal{D}}(\bar{r}, t)$, with units of coulombs per square meter (C/m^2), and

5. The *magnetic flux density,* $\bar{\mathcal{B}}(\bar{r}, t)$, with units of tesla (T), or webers per square meter (Wb/m^2).

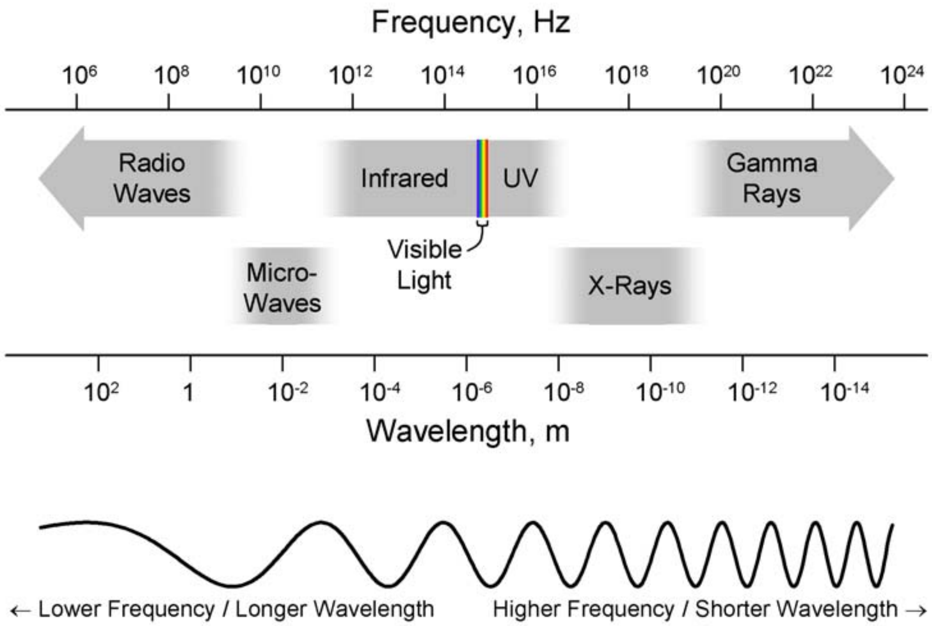


Figure 2.1: The electromagnetic spectrum.

Electromagnetic fields are created by stationary or moving charges. The stationary charges are described by the *electric charge density,* $\rho_v(\bar{r}, t)$, with units of coulombs per cubic meter (C/m^3). The *electric current densityi*s described by $\bar{\mathcal{J}}(\bar{r}, t)$ also describe moving charges, for example that exist in an antenna. In this chapter, we study the properties of electromagnetic waves without concerning ourselves with how the waves were created.

We will also see that the various fields and fluxes are related by properties of the materials in which the fields are propagating. The electric field intensity, $\bar{\mathcal{E}}(\bar{r}, t)$, and the electric flux density, $\bar{\mathcal{D}}(\bar{r}, t)$, are related by the electric permittivity, ε, with units of farads per meter (F/m). The magnetic field intensity, $\bar{\mathcal{H}}(\bar{r}, t)$, and the magnetic flux density, $\bar{\mathcal{B}}(\bar{r}, t)$, are related by the magnetic permeability, μ, with units of henries per meter (H/m). The electric current density $\bar{\mathcal{J}}(\bar{r}, t)$ is related to the electric field intensity, $\bar{\mathcal{E}}(\bar{r}, t)$ by the conductivity σ.

The parameters ε, μ and σ, referred to collectively as the constitutive parameters are in general functions of frequency, and also may be functions of space, direction, time, and even field intensity.

2.3 Maxwell's Equations

A vector field is completely defined, within an additive constant, by specifying the curl and the divergence of the field. The electromagnetic vector fields are described by curl and divergence relationships that are collectively known as Maxwell's equations. Maxwell's curl relationships are

$$\nabla \times \bar{\mathcal{E}}(\bar{r}, t) = -\frac{\partial \bar{\mathcal{B}}(\bar{r}, t)}{\partial t} \tag{2.2}$$

$$\nabla \times \bar{\mathcal{H}}(\bar{r}, t) = \frac{\partial \bar{\mathcal{D}}(\bar{r}, t)}{\partial t} + \bar{\mathcal{J}}(\bar{r}, t) \tag{2.3}$$

which are Faraday's law and Ampere's law, respectively. The term $\partial \bar{\mathcal{D}}(\bar{r},t)/\partial t$ is the displacement current present, for example, in a capacitor. For this reason, the electric flux density is also known as the *electric displacement*. The term $\bar{\mathcal{J}}(\bar{r},t)$ is the conduction current (such as the current in a wire) or the convection current (such as a stream of electrons in an oscilloscope).

Maxwell's divergence relationships are

$$\nabla \cdot \bar{\mathcal{D}}(\bar{r},t) = \rho_v(\bar{r},t) \tag{2.4}$$

$$\nabla \cdot \bar{\mathcal{B}}(\bar{r},t) = 0 \tag{2.5}$$

These two equations are sometimes referred to as Gauss' laws. For the sake of a simpler notation, the time and space dependence are often left off when writing the symbols for the various fields, currents, and charges. For example, the electric field $\bar{\mathcal{E}}(\bar{r},t)$ is written simply as $\bar{\mathcal{E}}$, the magnetic field $\bar{\mathcal{H}}(\bar{r},t)$ is written as $\bar{\mathcal{H}}$, etc.

Using the vector identity $\nabla \cdot (\nabla \times \bar{A}) = 0$ it follows that

$$\nabla \cdot \left(\nabla \times \bar{\mathcal{H}}\right) = \nabla \cdot \left(\frac{\partial \bar{\mathcal{D}}}{\partial t} + \bar{\mathcal{J}}\right) = 0 \tag{2.6}$$

By interchanging the order of the time and space derivatives operating on $\bar{\mathcal{D}}$ and using Gauss' law $\nabla \cdot \bar{\mathcal{D}} = \rho_v$, this can be rewritten as

$$\frac{\partial \rho_v}{\partial t} + \nabla \cdot \bar{\mathcal{J}} = 0 \tag{2.7}$$

This equation is the *continuity equation for the conservation of charge and current* and tells us that charges which vary in time give rise to currents, and that ρ_v and $\bar{\mathcal{J}}$ are dependent upon one another.

Lorentz force law relates the electromagnetic fields to measurable forces. It quantifies the force on a positive charge moving though electric and magnetic fields. Lorentz force law is written as

$$\bar{f} = q\bar{\mathcal{E}} + q\bar{v} \times \bar{\mathcal{B}} \tag{2.8}$$

where q is the charge (coulombs), $\bar{v}$ is the velocity (m/s), and $\bar{f}$ is the force (newtons). If the charge is not discrete, but rather given by a charge density ρ_v, then the law is modified to relate the electromagnetic fields to a *force density,* $\bar{\mathcal{F}}$, with units of N/m^3:

$$\bar{\mathcal{F}} = \rho_v\bar{\mathcal{E}} + \bar{\mathcal{J}} \times \bar{\mathcal{B}} \tag{2.9}$$

2.4 Wave Propagation in Free Space

We will first focus on the propagation of waves in free space (vacuum), because it is certainly quite astonishing that electromagnetic waves can propagate in a vacuum. How does the propagation of an electromagnetic wave through "nothing" occur? As stated previously, electromagnetic material properties relate the field quantities $\bar{\mathcal{E}}$, $\bar{\mathcal{H}}$, $\bar{\mathcal{D}}$ and $\bar{\mathcal{B}}$. These material properties are described by the constitutive parameters which, in a vacuum, are simply scalar constants. For a wave in free space, the constitutive relationships are

$$\bar{\mathcal{D}} = \varepsilon_o\bar{\mathcal{E}} \tag{2.10}$$

$$\bar{\mathcal{B}} = \mu_o\bar{\mathcal{H}} \tag{2.11}$$

where $\varepsilon_o = 8.8542 \times 10^{-12}$ F/m is the permittivity of free space, and $\mu_o = 4\pi \times 10^{-7}$ H/m is the permeability of free space. In free space the waves propagate without loss, and the conductivity is zero. The reason that these parameters are scalars is because free space is:

1. *Time-invariant:* the material properties do not vary with time.

2. *Homogeneous:* the material properties are constant throughout space.
3. *Isotropic:* the material properties are the same in all directions.
4. *Linear:* the electric and magnetic flux densities are proportional to the electric and magnetic fields, respectively. In other words, the material properties are independent of the field amplitudes.

In Chapter 6 we will study how time-varying electromagnetic fields are generated. For now, we are interested in the propagation of electromagnetic fields after they have been generated. In other words, we are interested in the propagation of fields in a region free of sources. If there are no sources of electromagnetic fields within the region of interest, then $\rho_v(\bar{r}, t) = 0$ and $\bar{\mathcal{J}}(\bar{r}, t) = 0$. In the absence of sources, Maxwell's equations in free space reduce to

$$\nabla \times \bar{\mathcal{E}}(\bar{r}, t) = -\mu_o \frac{\partial \bar{\mathcal{H}}(\bar{r}, t)}{\partial t} \tag{2.12}$$

$$\nabla \times \bar{\mathcal{H}}(\bar{r}, t) = \varepsilon_o \frac{\partial \bar{\mathcal{E}}\,(\bar{r}, t)}{\partial t} \tag{2.13}$$

$$\nabla \cdot \bar{\mathcal{E}}(\bar{r}, t) = 0 \tag{2.14}$$

$$\nabla \cdot \bar{\mathcal{H}}(\bar{r}, t) = 0 \tag{2.15}$$

Notice that $\nabla \cdot \bar{\mathcal{D}} = \nabla \cdot [\varepsilon_o \bar{\mathcal{E}}] = 0$, and since ε_o describes a homogeneous and isotropic region, Equation 2.14 follows. In a similar manner Equation 2.15 follows because μ_o does not depend on space.

The two curl equations (Equation 2.12 and Equation 2.13) can be manipulated to generate equations that contain only the electric field $\bar{\mathcal{E}}$ or the magnetic field $\bar{\mathcal{H}}$. For example, taking the curl of both sides of Equation 2.12 and using the vector identity $\nabla \times (\nabla \times \bar{A}) = \nabla(\nabla \cdot \bar{A}) - \nabla^2 \bar{A}$ results in the following

$$\nabla \times \left(\nabla \times \bar{\mathcal{E}}\right) = \nabla \times \left(-\mu_o \frac{\partial \bar{\mathcal{H}}}{\partial t}\right) = -\nabla^2 \bar{\mathcal{E}} \tag{2.16}$$

To derive this relationship we have used $\nabla(\nabla \cdot \bar{\mathcal{E}}) = 0$ which follows from Equation 2.14. Keep in mind this relationship is valid only in isotropic, homogeneous media. The operator ∇^2 is referred to as the *three-dimensional Laplacian operator*. In a Cartesian coordinate system this operator is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.17}$$

It is important to note that ∇^2 is a scalar operator, not a vector operator! Evaluating the middle term of Equation 2.16 yields

$$\nabla \times \left(-\mu_o \frac{\partial \bar{\mathcal{H}}}{\partial t}\right) = -\mu_o \frac{\partial}{\partial t}\left(\nabla \times \bar{\mathcal{H}}\right) = -\mu_o \frac{\partial}{\partial t}\left(\varepsilon_o \frac{\partial \bar{\mathcal{E}}}{\partial t}\right) \tag{2.18}$$

where Equation 2.13 was used to eliminate $\bar{\mathcal{H}}$. It then follows that:

$$\nabla^2 \bar{\mathcal{E}} - \mu_o \varepsilon_o \frac{\partial^2 \bar{\mathcal{E}}}{\partial t^2} = 0 \tag{2.19}$$

This equation is referred to as the *wave equation* for the electric field. Notice that Equation 2.19 contains only the electric field and the parameters ε_o and μ_o. The electric field wave equation can also be written as

$$\nabla^2 \bar{\mathcal{E}} - \frac{1}{c^2} \frac{\partial^2 \bar{\mathcal{E}}}{\partial t^2} = 0 \tag{2.20}$$

where

$$c = \frac{1}{\sqrt{\mu_o \varepsilon_o}} \approx 2.998 \times 10^8 \text{m/s} \tag{2.21}$$

is the speed of light in a vacuum.

An analogous derivation gives the *wave equation for the magnetic field*:

$$\nabla^2 \overline{\mathcal{H}} - \mu_o \varepsilon_o \frac{\partial^2 \overline{\mathcal{H}}}{\partial t^2} = 0 \tag{2.22}$$

A subtle and extremely important realization is that *not all solutions of the wave equations will satisfy Maxwell's equations*. In other words, just because a certain solution to the wave equation exists does not mean that it is a physically realizable electromagnetic field! The simple reason is because, when deriving the wave equations from Maxwell's equations, we have used $\nabla \cdot \overline{\mathcal{E}} = 0$ and $\nabla \cdot \overline{\mathcal{H}} = 0$ in order to eliminate $\nabla \cdot \overline{\mathcal{E}}$ and $\nabla \cdot \overline{\mathcal{H}}$ from the final form of the equations. Because of this, the requirement that the divergence of the electric or magnetic field is zero is no longer present in the wave equations! Therefore, in addition to their respective wave equations, an electric field of a wave *must satisfy* $\nabla \cdot \overline{\mathcal{E}} = 0$, and a magnetic field of a wave *must satisfy* $\nabla \cdot \overline{\mathcal{H}} = 0$.

What does the wave equation tell us about the properties of an electromagnetic wave? Taking the electric field as an example, the most general solution to the wave equation is

$$\overline{\mathcal{E}}(\overline{r}, t) = \overline{\mathcal{E}}_1 f_1(\omega t - \overline{k} \cdot \overline{r}) + \overline{\mathcal{E}}_2 f_2(\omega t + \overline{k} \cdot \overline{r}) \tag{2.23}$$

The first portion, $\overline{\mathcal{E}}_1 f_1(\omega t - \overline{k} \cdot \overline{r})$, is a wave that propagates *forward* in space as time increases, whereas the second portion, $\overline{\mathcal{E}}_2 f_2(\omega t + \overline{k} \cdot \overline{r})$, is a wave that propagates *backward* in space as time increases. The functions f_1 and f_2 can be any functions that propagate in space and time. The constant ω is the angular frequency with respect to time, and has units of radians/sec. The constant vector $\overline{k} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z}$ is referred to as the *wave vector* or *k-vector*, and its direction determines the direction of wave propagation. The magnitude of the wave vector, which is represented simply as k, is called the *wave number* or *propagation constant*, and can be viewed as a "'spatial frequency"', since it has units of radians/m. Note that the frequency in time and "'frequency in space" are not independent. In free space, we will soon show that $k_x^2 + k_y^2 + k_z^2 = \omega^2 \mu_o \varepsilon_o$, which is compactly written as

$$\overline{k} \cdot \overline{k} = k^2 = \omega^2 \mu_o \varepsilon_o = \frac{\omega^2}{c^2} \tag{2.24}$$

Equation 2.24 indicates that the magnitude of the wave vector in free space must always lie on a sphere of radius $\omega \sqrt{\mu_o \varepsilon_o}$ (see Chapter 1). This surface is called a "*k-vector surface*" and is an important concept in the course. In free space, $\overline{k}$ is a *purely real vector* and the k_x, k_y, and k_z components can be expressed in terms of angles as shown in Figure 2.2. We will see later that the wavevector can be complex in certian situations. Because of the importance of the wave vector, we summarize the concept:

Wave Vector Summary:

The vector $\overline{k} = k_x \hat{x} + k_y \hat{y} + k_z \hat{z}$ points in the direction of propagation and is called the wave vector. The magnitude of the wave vector is called the wave number or propagation constant, and is represented as $k = \sqrt{k_x^2 + k_y^2 + k_z^2} = \omega \sqrt{\mu_o \varepsilon_o}$ in free space. The propagation constant represents a "spatial frequency" since it has units of rad/m.

One of the simplest electromagnetic field solutions in free space is the sinusoidal wave

$$\overline{\mathcal{E}}(z, t) = \hat{x} E_o \cos(\omega t - kz) \tag{2.25}$$

which is plotted in Figure 2.3. This sinusoid has an amplitude in the x-direction and propagates in the z-direction in space as time increases. In the figure, the time- and space-dependence have been plotted on

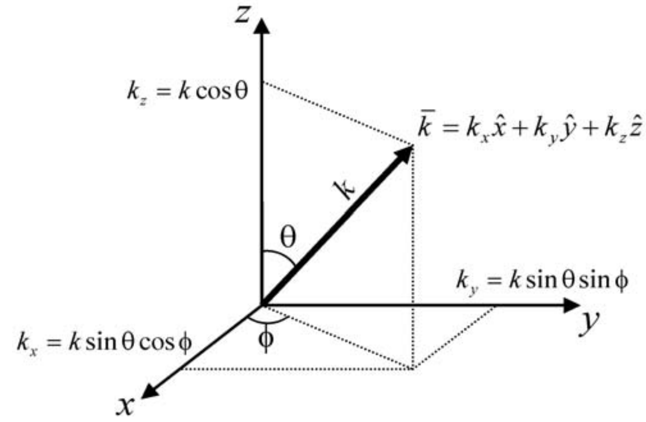


Figure 2.2: Description of the wave vector in the Cartesian and spherical coordinate systems. (Only when $\bar{k}$ is real can the vector be described in terms of angles.)

separate axes. To determine the direction of propagation, start with the wave at $t = 0$. Then, look at what happens to the wave as you increase time (progress to the right on the graph). Can you see the direction the wave is moving? This wave is a sinusoid that is propagating in the +z-direction. The period in time, T, can be determined by finding the time increment such that the amplitude of the wave at time $t = t_o$ and time $t = t_o + T$ are equal. In other words, we must find T such that

$$\bar{\mathcal{E}}\,(z_o, t_o) = \bar{\mathcal{E}}(z_o, t_o + T) \tag{2.26}$$

Evaluating this expression at $z_o = 0$ and $t_o = 0$ results in $1 = E_o \cos(\omega T)$. Discarding the trivial solution $T = 0$, the next possible value of ωT that satisfies this equation is $\omega T = 2\pi$, or

$$T = 2\pi/\omega = 1/f \tag{2.27}$$

The period in space (or wavelength), λ, is found in a similar manner. In this case, we must find the space increment such that the amplitude of the wave at positions z_o and $z_o + \lambda$ are equal, i.e. where

$$\bar{\mathcal{E}}(z_o, t_o) = \bar{\mathcal{E}}(z_o + \lambda, t_o) \tag{2.28}$$

Evaluating the expression at $z_o = 0$ and $t_o = 0$ leads to $1 = E_o \cos(k\lambda)$. Again, assuming that $k\lambda$ is nonzero, the next possible value that satisfies this equation is $k\lambda = 2\pi$, or

$$\lambda = 2\pi/k \tag{2.29}$$

This is an important relationship that relates the wavelength of an electromagnetic wave to the propagation constant, k.

Another simple solution to the wave equation is $\bar{\mathcal{E}}(z, t) = \hat{x}E_o \cos(\omega t + kz)$, which is a sinusoid that is propagating in the negative z-direction. This wave is illustrated in the Figure 2.4.

The velocity of the wave is obtained by "surfing" the wave, which means staying at one point on the wave as it moves. Mathematically speaking, this means that we want to find the velocity at which the amplitude of the wave does not change, i.e. the velocity such that $\omega t - kz = A$ or $z = (\omega t - A)/k$, where A is a constant. To find that velocity, we simply take the derivative of the position, z, with respect to time:

$$\frac{dz}{dt} = \frac{\omega}{k} = \frac{\omega}{\omega\sqrt{\mu_o\varepsilon_o}} = \frac{1}{\sqrt{\mu_o\varepsilon_o}} = c \approx 2.998 \times 10^8 \quad m/s \tag{2.30}$$

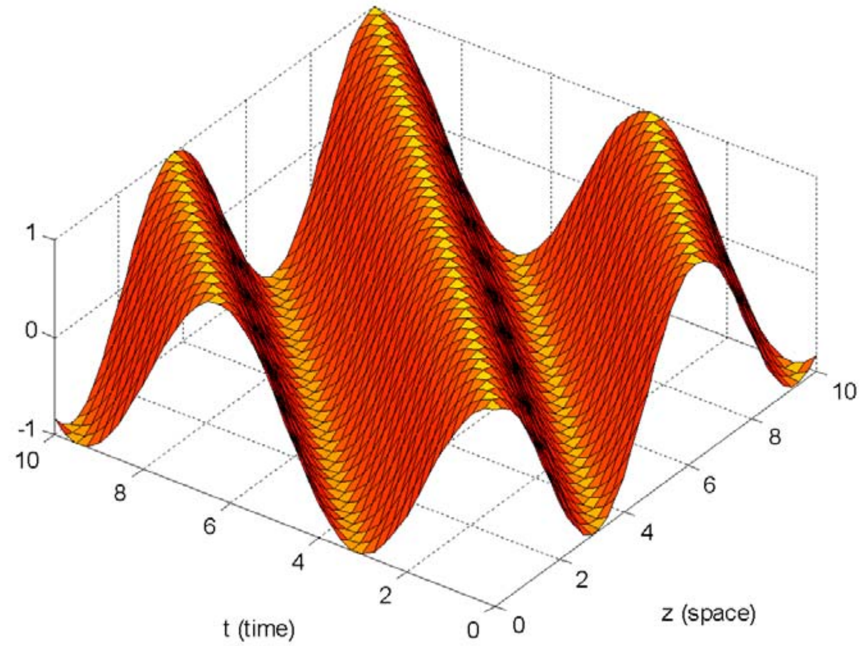


Figure 2.3: Plot of a sinusoid propagating in the positive z-direction as a function of space and time.

One of the triumphs of Maxwell's equations was the correct prediction of the speed of light well before it was precisely measured! In fact, it played an important role in making scientists of the time take note of Maxwell's theory, since it was unusual for a mathematical theory to *predict* experimental results!

Given the electric field solution $\bar{\mathcal{E}}(z,t) = \hat{x}E_o\cos(\omega t - kz)$, the magnetic field is calculated from Maxwell's equations. Using Faraday's law in free space (Equation 2.12), we find that

$$\frac{\partial \bar{\mathcal{H}}}{\partial t} = -\frac{1}{\mu_o}\nabla\times\bar{\mathcal{E}} = -\frac{1}{\mu_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_o\cos(\omega t - kz) & 0 & 0 \end{vmatrix}$$

$$= -\hat{y}\frac{kE_o}{\mu_o}\sin(\omega t - kz) \tag{2.31}$$

To find $\bar{\mathcal{H}}$, we simply integrate this over time:

$$\bar{\mathcal{H}} = -\hat{y}\frac{kE_o}{\mu_o}\int \sin(\omega t - kz)\, dt = \hat{y}\frac{kE_o}{\omega\mu_o}\cos(\omega t - kz) \tag{2.32}$$

Given this result, you may wonder what happened to the constant of integration? The simple answer is, we are only interested in time-varying fields, and therefore the constant is of no interest to us. Using the fact that the free space propagation constant is $k = \omega\sqrt{\mu_o\varepsilon_o}$, the magnetic field solution can be written as

$$\bar{\mathcal{H}}(z,t) = \hat{y}\sqrt{\frac{\varepsilon_o}{\mu_o}}E_o\cos(\omega t - kz) \tag{2.33}$$

From this solution, the following observations are made:

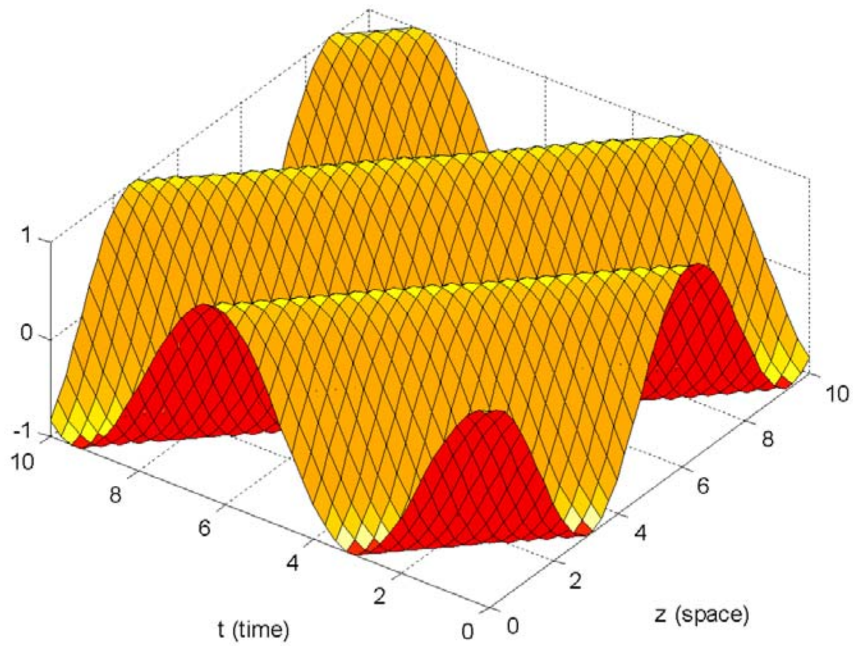


Figure 2.4: Plot of a sinusoid propagating in the negative z-direction as a function of space and time.

1. Both the electric and magnetic fields propagate at a velocity of $c = 1/\sqrt{\mu_o \varepsilon_o}$ in free space.
2. The magnetic and electric fields are orthogonal in space.
3. The plane that is formed by the electric and magnetic fields (the x-y plane in this example) is perpendicular to the direction of propagation (the z-direction).
4. The magnetic field solution has the same functional dependence as the electric field (i.e. $\sin(\omega t - kz)$ in this example), but differ in amplitude by a factor of $\sqrt{\varepsilon_o/\mu_o}$. The *inverse* of this factor is referred to as the *wave impedance of free space*, η_o, and has units of Ohms (Ω). In free space, the wave impedance is:

$$\eta_o = \sqrt{\frac{\mu_o}{\varepsilon_o}} \approx 377 \ \ \Omega \tag{2.34}$$

These observations are made for a specific case, so the next logical question is: are these observations true in general? As we will see in the following section, they are.

2.5 Plane Waves in Free Space

Maxwell's equations are linear. For this reason, electromagnetic fields that have a general functional dependence in time can be described by a superposition of time-harmonic fields using Fourier analysis. These harmonic waves can be written in their sinusoidal, time-domain form, but such functions are cumbersome to manipulate. The analysis is simplified by defining a complex vector phasor, which is an extension of the phasors used in circuit theory. The phasors used in electromagnetics are more general since they capture *both spatial and time dependence.* In addition to superposition in time, superposition in space is also used to analyze the propagating characteristics of arbitrary wave fronts (for example, waves that radiate out spherically from a point source). The basic "harmonic" spatial component is the plane wave. The reason why the wave is called a plane wave will soon become clear.

2.5.1 Vector Phasors

First, let us see how the spatial component enters into the phasor. For a simple cosine wave:

$$\bar{\mathcal{E}}\,(z,t) = \hat{x}E_{\text{xo}}\cos(\omega t - k_z z) = \mathcal{R}e\left\{E_{\text{xo}}e^{-jk_z z}e^{j\omega t}\right\}\hat{x} \tag{2.35}$$

where E_{xo} is the amplitude of the wave. Note how the spatial and time dependence separates into two multiplicative terms. Assuming that the time dependence is known, we are only interested in the amplitude, spatial dependence, and the resulting phase shifts. We therefore define the phasor E_x as

$$E_x = E_{\text{xo}}e^{-jk_z z} \tag{2.36}$$

such that the wave can be written as $\bar{\mathcal{E}}(z,t) = \mathcal{R}e\{E_x e^{j\omega t}\}\hat{x}$.

Up to this point, we have assumed that the electric field and the propagation direction coincide with the axes of a coordinate system. Generalizing the result, a general phasor for waves propagating in an arbitrary direction is

$$\begin{aligned}\bar{\mathcal{E}}(\bar{r},t) &= \mathcal{R}e\left\{E_{\text{xo}}e^{j(\omega t - k_x x - k_y y - k_z z)}\hat{x} + E_{\text{yo}}e^{j(\omega t - k_x x - k_y y - k_z z)}\hat{y} + E_{\text{zo}}e^{j(\omega t - k_x x - k_y y - k_z z)}\hat{z}\right\}\\ &= \mathcal{R}e\left\{\bar{E}e^{j\omega t}\right\}\end{aligned} \tag{2.37}$$

where we now define the *vector phasor*

$$\bar{E} = E_{\text{xo}}e^{-j\bar{k}\cdot\bar{r}}\hat{x} + E_{\text{yo}}e^{-j\bar{k}\cdot\bar{r}}\hat{y} + E_{\text{zo}}e^{-j\bar{k}\cdot\bar{r}}\hat{z} = \bar{E}_o e^{-j\bar{k}\cdot\bar{r}} \tag{2.38}$$

Therefore, for waves propagating in an arbitrary direction, the phasor is a *complex-valued vector* which is a function of position. Since the time-dependence and the direction of propagation are known, we are only interested in the amplitude, spatial dependence, and the resulting phase shifts of each of the phasor components E_{xo}, E_{yo}, and E_{zo}.

As with scalar phasors in circuit theory, the basic idea behind using phasors in electromagnetics is to manipulate the complex-valued vector phasor, and then to return to the "real world" by taking the real part after multiplying by $e^{j\omega t}$. In electromagnetics, however, we rarely go through this whole process since we will learn to interpret the phasor quantities directly. Because of the importance of vector phasors in electromagnetics, we summarize their properties:

Vector Phasor Summary:

The phasor $\bar{E}$ is a complex vector quantity. In its most general form, $\bar{E} = \bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$, where $\bar{r} = x\hat{x} + y\hat{y} + z\hat{z}$ is the position vector, and $\bar{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$ is the wave vector. The vector quantity $\bar{E}_o = E_{\text{xo}}\hat{x} + E_{\text{yo}}\hat{y} + E_{\text{zo}}\hat{z}$ is a constant vector that represents the initial amplitude of the wave, and may be complex. The phasor can also be written as $\bar{E} = E_x\hat{x} + E_y\hat{y} + E_z\hat{z}$, where $E_x = E_{\text{xo}}e^{-j\bar{k}\cdot\bar{r}}$, $E_y = E_{\text{yo}}e^{-j\bar{k}\cdot\bar{r}}$, and $E_z = E_{\text{zo}}e^{-j\bar{k}\cdot\bar{r}}$. The time-domain representation of the wave is reconstructed by the transformation $\bar{\mathcal{E}}(\bar{r},t) = \mathcal{R}e\{\bar{E}e^{j\omega t}\}$.

We can convert Maxwell's equations to phasor form by substituting the phasor fields $\bar{\mathcal{E}}(\bar{r},t) = \mathcal{R}e\{\bar{E}e^{j\omega t}\}$, $\bar{\mathcal{B}}(\bar{r},t) = \mathcal{R}e\{\bar{B}e^{j\omega t}\}$, $\bar{\mathcal{H}}(\bar{r},t) = \mathcal{R}e\{\bar{H}e^{j\omega t}\}$, and so forth. The result is that derivatives with respect to time become a simple multiplication by $j\omega$. The resulting equations are referred to as the *time-harmonic form of Maxwell's equations* (see Appendix B for the derivation of this relationship)

$$\nabla\times\bar{E} = -j\omega\bar{B} \tag{2.39}$$

$$\nabla \times \bar{H} = j\omega\bar{D} + \bar{J} \tag{2.40}$$

$$\nabla \cdot \bar{D} = \rho_v \tag{2.41}$$

$$\nabla \cdot \bar{B} = 0 \tag{2.42}$$

In *free space* without any sources present, Maxwell's equations in time-harmonic form become:

$$\nabla \times \bar{E} = -j\omega\mu_o\bar{H} \tag{2.43}$$

$$\nabla \times \bar{H} = j\omega\varepsilon_o\bar{E} \tag{2.44}$$

$$\nabla \cdot \bar{E} = 0 \tag{2.45}$$

$$\nabla \cdot \bar{H} = 0 \tag{2.46}$$

For time-harmonic fields, the electric field *wave equation* may also be written in the time-harmonic form

$$\nabla^2\bar{E} + \omega^2\mu_o\varepsilon_o\bar{E} = 0$$

which is commonly written as

$$\nabla^2\bar{E} + k^2\bar{E} = 0 \tag{2.47}$$

This is known as the *Helmholtz equation.* The general solution of Helmholtz equation is (see Appendix B):

$$\bar{E} = \bar{E}_+ e^{-j\bar{k}\cdot\bar{r}} + \bar{E}_- e^{j\bar{k}\cdot\bar{r}} \tag{2.48}$$

where the vector quantities $\bar{E}_+$ and $\bar{E}_-$ are the amplitudes of forward and backward travelling waves, respectively. In addition, recall that the wave vector must satisfy

$$\bar{k} \cdot \bar{k} = \omega^2\mu_o\varepsilon_o \tag{2.49}$$

in free space, where $|\bar{k}| = k = \omega\sqrt{\mu_o\varepsilon_o} = \omega/c = 2\pi/\lambda$, and λ is the wavelength in free space. As mentioned previously, this relationship defines *a sphere in k-space* with a radius that depends on the frequency and the electromagnetic properties of the medium. *This means that no matter which direction the wave is propagating, the wave vector of the plane wave will always have the same magnitude.*

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 2.1:

An electromagnetic wave propagating in free space has an electric field vector given by

$$\bar{E} = j20e^{-j0.01x+j0.03z}\hat{y}$$

Find the value of the constant vector $\bar{E}_o$, the wave vector $\bar{k}$, and the frequency of the wave f (in Hz).

Solution:

Remember that the general phasor form of an electromagnetic wave is

$$\bar{E} = \bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$$

By inspection, we see that $\bar{E}_o = j20\hat{y}$. Also, the wave vector must be $\bar{k} = 0.01\hat{x} - 0.03\hat{z}$ (carry out the dot product to see this for yourself). The frequency of the wave can be found from the magnitude of the wave vector:

$$k = \frac{\omega}{c} = \frac{2\pi f}{c}$$

Solving for frequency:

$$f = k\frac{c}{2\pi} = \sqrt{0.01^2 + 0.03^2}\left(\frac{3\times 10^8}{2\pi}\right) = 1.51\,\text{MHz}$$

$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

EXAMPLE 2.2:

Find the time-domain form of the electric field, $\bar{\mathcal{E}}$, from the previous example. Also, find the phasor form of the corresponding magnetic field, $\bar{H}$.

Solution:

The time-domain representation of the electric field is reconstructed by the transformation

$$\bar{\mathcal{E}} = \mathcal{R}e\left\{\bar{E}e^{j\omega t}\right\} = \mathcal{R}e\left\{\left(j20e^{-j0.01x+j0.03z}\hat{y}\right)e^{j\omega t}\right\} = \mathcal{R}e\left\{j20e^{j(\omega t-0.01x+0.03z)}\hat{y}\right\}$$

$$= \mathcal{R}e\left\{j20\left[\cos(\omega t - 0.01x + 0.03z) + j\sin(\omega t - 0.01x + 0.03z)\right]\hat{y}\right\}$$

$$= \mathcal{R}e\left\{j20\,\cos(\omega t - 0.01x + 0.03z) - 20\sin(\omega t - 0.01x + 0.03z)]\,\hat{y}\right\}$$

$$= -20\sin(\omega t - 0.01x + 0.03z)]\,\hat{y}$$

The phasor form of the magnetic field is found using the phasor form of Faraday's law in free space:

$$\bar{H} = \frac{j}{\omega\mu_o}\nabla\times\bar{E} = \frac{j}{\omega\mu_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & j20e^{-j0.01x+j0.03z} & 0 \end{vmatrix}$$

$$= \frac{j}{\omega\mu_o}[-\hat{x}(j0.03)j20e^{-j0.01x+j0.03z} + \hat{z}(-j0.01)j20e^{-j0.01x+j0.03z}]$$

$$= \frac{j}{\omega\mu_o}\left(0.6\hat{x} + 0.2\hat{z}\right)e^{-j0.01x+j0.03z}$$

Recall that the permeability of free space is $\mu_o = 4\pi\times 10^{-7}$ H/m. Also, from the previous example we have $\omega = 2\pi\times 1.51\times 10^6 = 9.49\times 10^6$ rad/s. Substituting these values we find

$$\bar{H} = j\left(7.15\hat{x} + 2.38\hat{z}\right)e^{-j0.01x+j0.03z}$$

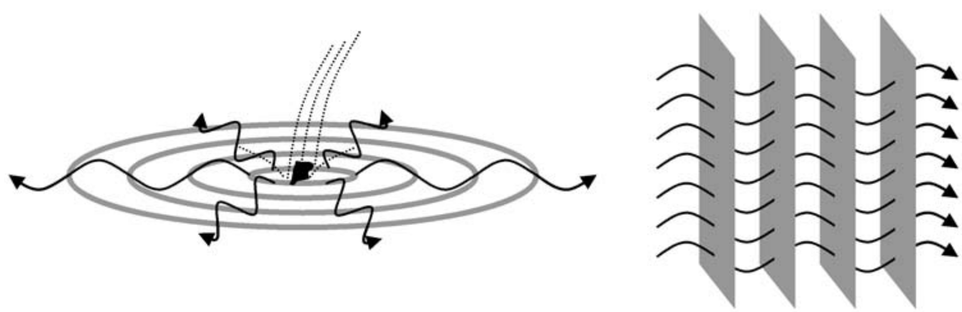
$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

2.5.2 Why is a Plane Wave Called a "Plane" Wave?

The "type" of a wave is described by the geometric shape of the locations in space at which the phase is constant at any time. For example, think of dropping a stone into water. A wave propagates in concentric circles away from the disturbance. The troughs or peaks that propagate radially outward represent the places at which the phase is constant. Because the wave only propagates in two dimensions (on the water's surface), the wave exhibits "circles of constant phase," as shown in Figure 2.5(a). Therefore, we could call this wave a "circular wave."

Electromagnetic waves propagate in three-dimensional space, and the places at which the phase is a constant at any time form a three-dimensional surface. For this reason, we describe waves propagating in space as "plane" waves, "spherical" waves, "cylindrical" waves, etc. For example, if an electromagnetic wave were propagating outward in all directions from a point source, the surfaces of constant phase would be three-dimensional spheres, and this wave would be called a "spherical wave."

For a wave of the form $\bar{E} = \bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$ propagating in free space, the wave vector $\bar{k}$ has a fixed magnitude $k = \omega\sqrt{\mu_o \varepsilon_o}$, which is determined only by the material in which the wave is propagating. The position vector $\bar{r} = x\hat{x} + y\hat{y} + z\hat{z}$ varies throughout space. The surface of constant phase, then, is defined as any surface in which the phase of the wave is a constant, i.e. where $\bar{k}\cdot\bar{r}$ is a constant. Recall that the vector relationship $\bar{k}\cdot\bar{r} = d$ describes a plane (see Chapter 1). Therefore, the physical interpretation of $\bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$ is that it describes a *plane wave*, as shown in Figure 2.5(b).



(a) (b)

Figure 2.5: (a) When a stone is dropped in water, waves propagate radially outward on the water's surface. This wave could be called a "circular wave" since the peaks and troughs form circles of constant phase. (b) For an electromagnetic wave of the form $\bar{E} = \bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$ propagating in three-dimensional space, the peaks and troughs form planes of constant phase, and therefore this wave is called a "plane wave."

2.5.3 Properties of Plane Waves in Free Space

From the free space Maxwell's equations in time-harmonic form, we can establish the relationships between the electromagnetic fields and the wave vector, $\bar{k}$. The spatial relationship between the electric field vector, $\bar{E}$, and $\bar{k}$ can be determined by first noting that $\nabla\cdot\bar{E} = -j\bar{k}\cdot\bar{E}$. Then, Maxwell's divergence equation for the electric field becomes $-j\bar{k}\cdot\bar{E} = 0$ or more simply,

$$\bar{k}\cdot\bar{E} = 0 \tag{2.50}$$

(refer to Appendix B). Recall that two vectors are orthogonal if their dot product is zero. This means the electric field vector, $\bar{E}$, is *perpendicular* to the direction of propagation, $\bar{k}$. In a similar manner, Maxwell's divergence equation for the magnetic field, $\nabla \cdot \bar{H} = 0$, can be manipulated to yield

$$\bar{k} \cdot \bar{H} = 0 \tag{2.51}$$

This means the magnetic field vector, $\bar{H}$, is *also perpendicular* to the direction of propagation, $\bar{k}$!

Finally, what is the relationship between the electric and magnetic fields? This relationship is found from Equation 2.43 which, after some manipulation, can be written as

$$\bar{H} = \frac{1}{\eta_o}\hat{k} \times \bar{E} \tag{2.52}$$

where $\eta_o = \sqrt{\mu_o/\varepsilon_o}$ is the free space wave impedance, and $\hat{k}$ is a unit vector in the direction of propagation (a vector in the direction of $\bar{k}$ but with unity magnitude). Recall that the cross product of two vectors results in a vector that is orthogonal to *both vectors.* We have already shown that $\bar{E}$ is orthogonal to $\bar{k}$, so therefore the magnetic field is perpendicular to the electric field and the direction of propagation. This means that $\bar{k}$, $\bar{E}$ and $\bar{H}$ form a *right handed coordinate system,* which makes it very easy to determine the directions of $\bar{k}$, $\bar{E}$, or $\bar{H}$ once the other two are known! Also from Equation 2.52 we can see that, in free space, the electric and magnetic fields oscillate *in-phase* since η_o is a real quantity. The relationships between $\bar{k}$, $\bar{E}$ and $\bar{H}$ are illustrated in Figure 2.6.

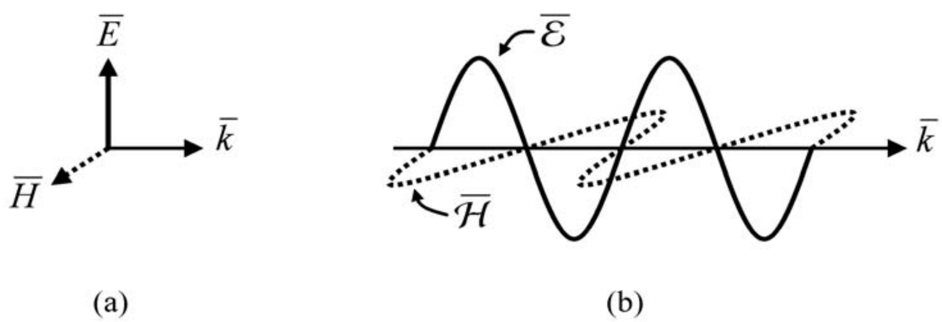


Figure 2.6: (a) In the phasor domain, the $\bar{k}$, $\bar{E}$ and $\bar{H}$ vectors form a right-handed coordinate system for wave propagation in free space. (b) The corresponding electric and magnetic fields in the time domain. Notice that the fields oscillate in-phase when propagating through free space.

A more direct way to come to the preceding conclusion is to start with the time-harmonic Maxwell's equations and rewrite them explicitly for plane waves. This has already been done for Maxwell's divergence equations, which were found to simplify to $\bar{k} \cdot \bar{E} = 0$ and $\bar{k} \cdot \bar{H} = 0$ for plane waves. Additionally, it can be shown that the curl operations simplify to $\nabla \times \bar{E} = -j\bar{k} \times \bar{E}$ and $\nabla \times \bar{H} = -j\bar{k} \times \bar{H}$ for the electric and magnetic fields, respectively. (To show this, write out the vector $\bar{E} = (E_{\text{xo}}\hat{x} + E_{\text{yo}}\hat{y} + E_{\text{zo}}\hat{z})e^{-j\bar{k}\cdot\bar{r}}$, and $\bar{k} \cdot \bar{r} = k_x x + k_y y + k_z z$. Then, explicitly carry out the cross product operation.) It then follows that Maxwell's equations for a *plane wave propagating in free space* are:

$$\bar{k} \times \bar{E} = \omega\mu_o\bar{H} \tag{2.53}$$

$$\bar{k} \times \bar{H} = -\omega\varepsilon_o\bar{E} \tag{2.54}$$

$$\bar{k} \cdot \bar{E} = 0 \tag{2.55}$$

$$\bar{k} \cdot \bar{H} = 0 \tag{2.56}$$

The orthogonality between the wave vector and each of the vector fields is clearly evident in these equations.

2.6 Wave Propagation in General Media

Free space (i.e. vacuum) is the simplest electromagnetic medium, and its permittivity and permeability can be expressed by the simple scalar constants ε_o and μ_o. In general, however, the electromagnetic properties of a material may depend upon time, position, the direction that the vector fields are oriented, and the intensity of the fields.

2.6.1 Characteristics of General Media

The properties of a medium may depend upon many factors. Four commonly considered factors are summarized next:

Time-variance

If the properties depend upon time, the material is called *time-variant.* An example of this is the air above a road as the road is heated by the sun. When travelling on a straight road on a sunny day, stationary objects in the distance appear to waver. This is caused by the electromagnetic properties of the air above the road changing with time due to the heat from the road. In this course, we will concern ourselves with materials that have *time-invariant* material properties, i.e. materials whose properties do not change over time.

Homogeneity

If the material properties are dependent upon the position within the material, the material is described as *inhomogeneous.* Materials that are inhomogeneous include glass or plastic that is graded to be darker on the top and lighter on the bottom, such as in car windshields or sunglass lenses. In both cases, a substance is added to the glass or plastic to change the material properties in a certain region. On the other hand, if the material properties are space-invariant, i.e. they do not depend on position, the material is described as *homogeneous.*

Isotropy

If the material properties are different depending upon the orientation of the fields propagating through the material, it is an *anisotropic* material. An example of this type of material is the lens from a pair of polarized sunglasses. Looking through the lens, the reflection from a lake may be clearly evident; however, by merely rotating the lens, the reflected light disappears. Another example is the liquid in a liquid crystal display (LCD). For light to be transmitted through the liquid layer, the electric field must be aligned with the molecules in the liquid crystal. When a voltage is applied, the molecules rotate and no longer transmit the electric field. (Note that a material that is *anisotropic* may still be *homogeneous*, if it exhibits the same anisotropy throughout the material.) If the orientation of the material with respect to the fields does not matter, then the material is *isotropic.*

Linearity

Finally, if the material properties are functions of the field intensity, the material is described as *nonlinear.* In this case, the fields and related fluxes do not have simple linear relationships. In this course, we will restrict our study to materials that are *linear*, which encompasses most materials.

For a material that is *linear*, *time-invariant*, *homogeneous*, and *isotropic*, the electric field and flux are related by

$$\bar{\mathcal{D}}(\bar{r},t) = \varepsilon\bar{\mathcal{E}}(\bar{r},t) \tag{2.57}$$

In this case, the permittivity is a simple scalar. In free space, $\varepsilon = \varepsilon_o$, as we have already seen. For other materials, we define $\varepsilon = \varepsilon_o\varepsilon_r$, where ε_r is called the *relative permittivity*, which is the permittivity of the material as compared to the permittivity of free space. In general, the permittivity of any material is larger than that of free space, so $\varepsilon > \varepsilon_o$ and $\varepsilon_r > 1$. In addition, in nonmagnetic media we define the *index of refraction* of a material as $n = \sqrt{\varepsilon_r}$. The index of refraction is most commonly used when working in the optical portion of the spectrum. Also, in general $n > 1$.

Now assume a general material that is still *linear*, but is *time-variant*, *inhomogeneous*, and *anisotropic*. In this case the permittivity is no longer a simple scalar quantity. Instead, the electric field and flux are related by

$$\bar{\mathcal{D}}(\bar{r},t) = \tilde{\varepsilon}(\bar{r},t)\bar{\mathcal{E}}(\bar{r},t) \tag{2.58}$$

where $\tilde{\varepsilon}(\bar{r},t)$ is the permittivity *tensor*. A tensor is matrix that describes a physical quantity. In a Cartesian coordinate system, this constitutive relationship becomes

$$\begin{pmatrix} \mathcal{D}_x \\ \mathcal{D}_y \\ \mathcal{D}_z \end{pmatrix} = \begin{pmatrix} \varepsilon_{\mathrm{xx}}(\bar{r},t) & \varepsilon_{\mathrm{xy}}(\bar{r},t) & \varepsilon_{\mathrm{xz}}(\bar{r},t) \\ \varepsilon_{\mathrm{yx}}(\bar{r},t) & \varepsilon_{\mathrm{yy}}(\bar{r},t) & \varepsilon_{\mathrm{yz}}(\bar{r},t) \\ \varepsilon_{\mathrm{zx}}(\bar{r},t) & \varepsilon_{\mathrm{zy}}(\bar{r},t) & \varepsilon_{\mathrm{zz}}(\bar{r},t) \end{pmatrix} \begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \\ \mathcal{E}_z \end{pmatrix} \tag{2.59}$$

Notice that the time-variant behavior is included in the variable t, and the inhomogeneous behavior by the position vector $\bar{r}$. The anisotropy causes the permittivity to be expressed as a matrix, and in this way captures the orientation dependence of the material. As before, we can define $\tilde{\varepsilon}(\bar{r},t) = \varepsilon_o\tilde{\varepsilon}_r(\bar{r},t)$, where $\tilde{\varepsilon}_r(\bar{r},t)$ is now called the *relative permittivity tensor*.

For a material that is *linear* and *time-invariant*, the relationship becomes

$$\begin{pmatrix} \mathcal{D}_x \\ \mathcal{D}_y \\ \mathcal{D}_z \end{pmatrix} = \begin{pmatrix} \varepsilon_{\mathrm{xx}}(\bar{r}) & \varepsilon_{\mathrm{xy}}(\bar{r}) & \varepsilon_{\mathrm{xz}}(\bar{r}) \\ \varepsilon_{\mathrm{yx}}(\bar{r}) & \varepsilon_{\mathrm{yy}}(\bar{r}) & \varepsilon_{\mathrm{yz}}(\bar{r}) \\ \varepsilon_{\mathrm{zx}}(\bar{r}) & \varepsilon_{\mathrm{zy}}(\bar{r}) & \varepsilon_{\mathrm{zz}}(\bar{r}) \end{pmatrix} \begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \\ \mathcal{E}_z \end{pmatrix} \tag{2.60}$$

and if the material is also *homogeneous*, the relationship further reduces to

$$\begin{pmatrix} \mathcal{D}_x \\ \mathcal{D}_y \\ \mathcal{D}_z \end{pmatrix} = \begin{pmatrix} \varepsilon_{\mathrm{xx}} & \varepsilon_{\mathrm{xy}} & \varepsilon_{\mathrm{xy}} \\ \varepsilon_{\mathrm{yx}} & \varepsilon_{\mathrm{yy}} & \varepsilon_{\mathrm{yz}} \\ \varepsilon_{\mathrm{zx}} & \varepsilon_{\mathrm{zy}} & \varepsilon_{\mathrm{zz}} \end{pmatrix} \begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \\ \mathcal{E}_z \end{pmatrix} \tag{2.61}$$

And finally, if the material is also *isotropic*, the relationship reduces to the familiar form where ε is simply a scalar quantity:

$$\begin{pmatrix} \mathcal{D}_x \\ \mathcal{D}_y \\ \mathcal{D}_z \end{pmatrix} = \varepsilon \begin{pmatrix} \mathcal{E}_x \\ \mathcal{E}_y \\ \mathcal{E}_z \end{pmatrix} = \varepsilon\bar{\mathcal{E}} = \varepsilon_r\varepsilon_o\bar{\mathcal{E}} \tag{2.62}$$

2.6.2 Polarizable and Magnetizable Media

What causes the permittivity of a material to be different than that of free space, and why in general is $\varepsilon > \varepsilon_o$? Electromagnetic radiation is created by oscillating charges. In a material, the electrons oscillating around a nuclei can be viewed as oscillators. When a wave propagates in a medium, the electromagnetic fields interact with the oscillators, and the electromagnetic fields produced by the oscillators in turn change the electromagnetic fields of the wave. In essence, the electromagnetic field "shakes" the collection of oscillators, and this interaction causes the wave to slow down. A model for this effect is that of a *polarizable medium*, which is a medium whose atomic charge distributions are altered by the presence of an electric field. We next examine how this interaction can be described mathematically. Since Maxwell's equations are applicable at a macroscopic scale, the material properties must be described at this level. To do this, however, we must first examine the material at a microscopic scale.

At a microscopic level, a polarizable medium is composed of atoms or molecules whose internal charge distributions are rearranged by the application of an electric field. While a complete model for these atoms and molecules is beyond the scope of this course, a simple model consists of viewing them as individual pairs of oscillating charges. These oscillators could be a proton and electron, a dipole in a organic molecule, or an ion-pair. A single oscillator can be thought of as two opposite charges, $+q$ and $-q$, separated by a distance $\bar{x}$, which is a time-harmonic quantity. The dipole moment for this pair is then $\bar{p} = q\bar{x}$. To obtain a macroscopic quantity, we average over a large number of oscillators and define a macroscopic polarization $\bar{\mathcal{P}} = Nq\bar{x}$, where N is the density of dipoles in the material. The quantity N has units of $1/\text{m}^3$, q has units of coulombs, and $\bar{x}$ is meters, and therefore $\bar{\mathcal{P}}$ has the same units as $\bar{\mathcal{D}}$, which is C/m^2. We then add $\bar{\mathcal{P}}$ to obtain the total electric flux density:

$$\bar{\mathcal{D}} = \varepsilon_o \bar{\mathcal{E}} + \bar{\mathcal{P}} \tag{2.63}$$

For a simple polarizable media, the macroscopic polarization is related to electric field intensity. Let us describe the dipole interaction by a mechanical oscillator consisting of a massive proton to which the electron is bound by a spring. We then equate the electrostatic force between the pair, $q\bar{\mathcal{E}}$, to the mechanical restoring force of the spring, $k\bar{x}$, where k is the spring constant. The value of the spring constant is related to the resonant frequency of the mechanical oscillation, ω_o, by $k = m_e\omega_o^2$, where m_e is the electron mass. We then solve for $\bar{x} = \bar{\mathcal{E}}q/k = \bar{\mathcal{E}}q/m_e\omega_o^2$. Substituting this back into the macroscopic polarization $\bar{\mathcal{P}} = Nq\bar{x}$, we find

$$\bar{\mathcal{P}} = \left(\frac{Nq^2}{m_e\omega_o^2}\right)\bar{\mathcal{E}} = \varepsilon_o\chi_e\bar{\mathcal{E}} \tag{2.64}$$

where χ_e is the *electric susceptibility*. From this result, we observe that the macroscopic polarization is linearly related to the electric field. Substituting this result for $\bar{\mathcal{P}}$ back into Equation 2.63, we obtain

$$\bar{\mathcal{D}} = \varepsilon_o\bar{\mathcal{E}} + \bar{\mathcal{P}} = \varepsilon_o(1+\chi_e)\bar{\mathcal{E}} = \varepsilon_o\varepsilon_r\bar{\mathcal{E}} = \varepsilon\bar{\mathcal{E}} \tag{2.65}$$

where ε is the *dielectric constant* or *electric permittivity* and ε_r is the *relative dielectric constant* or *relative electric permittivity*, as defined previously. The relative dielectric constant is typically between 1 and 100. In general, ε depends on the frequency, but more about this later.

A *magnetizable* medium is described by

$$\bar{\mathcal{B}} = \mu_o(\bar{\mathcal{H}} + \bar{\mathcal{M}}) \tag{2.66}$$

where $\bar{\mathcal{M}}$ is the macroscopic magnetization, which depends on the magnetic field intensity. In many media this relationship is nonlinear, for example in a magnet. If the relationship is linear, it can be expressed as $\bar{\mathcal{B}} = \mu\bar{\mathcal{H}}$. In this course, we will focus our attention on *nonmagnetic* media where $\bar{\mathcal{B}} = \mu_o\bar{\mathcal{H}}$.

The analysis of plane waves in free space is next extended to media other than free space that are *time-invariant*, *linear*, *homogeneous*, and *isotropic*. First, we will consider lossless and lossy media. Then we will consider propagation in a material where the electromagnetic properties depend on the frequency (or wavelength) of operation. Such a material is called a dispersive material.

2.6.3 Plane Waves in Lossless Media

For wave propagation in a lossless medium other than free space, the generalization of the permittivity to $\varepsilon = \varepsilon_o\varepsilon_r$ makes it simple to extend the concepts learned for free space wave propagation to other media. In a lossless nonmagnetic medium, the wave vector becomes

$$k = \omega\sqrt{\mu_o\varepsilon} = \omega\sqrt{\mu_o\varepsilon_o\varepsilon_r} = \sqrt{\varepsilon_r}\,(\omega\sqrt{\mu_o\varepsilon_o}) = nk_o \tag{2.67}$$

where we now define

$$k_o = \omega\sqrt{\mu_o\varepsilon_o} \tag{2.68}$$

as the *propagation constant in free space*, and

$$n = \frac{k}{k_o} = \sqrt{\varepsilon_r} \tag{2.69}$$

as the *index of refraction* of the material, as described earlier. This is an important conclusion, as it allows us to easily find the magnitude of the wave vector in any lossless material given the frequency or wavelength of the wave. In addition, the wave impedance in a medium decreases by a factor of $1/n$:

$$\eta = \sqrt{\frac{\mu_o}{\varepsilon_o \varepsilon_r}} = \frac{1}{n}\sqrt{\frac{\mu_o}{\varepsilon_o}} = \frac{\eta_o}{n} \tag{2.70}$$

where η_o is the *wave impedance of free space.* Assuming the material is isotropic, the spatial relationships between the field components remain unchanged (they remain orthogonal) and Equations 2.53-2.56 still apply, with $\overline{k} = nk_o\hat{k}$, where $\hat{k}$ is a unit vector in the direction of propagation. The velocity at which the plane wave propagates, v, is now slowed by a factor of $1/n$:

$$v = \frac{\omega}{k} = \frac{\omega}{nk_o} = \frac{c}{n} \tag{2.71}$$

which makes physical sense, since the velocity of a wave in a dense medium will be slower (think of trying to swim through molasses as opposed to water!). In free space, a plane wave propagates unobstructed, but in another medium the wave is slowed by the interaction with atoms, molecules and ions. Finally, the wavelength of a wave at a given frequency is less than the wavelength in free space by a factor of $1/n$:

$$\lambda = \frac{2\pi}{k} = \frac{2\pi}{nk_o} = \frac{\lambda_o}{n} \tag{2.72}$$

In summary, by introducing a relative permittivity, ε_r, we can easily understand the propagation of a plane wave in a lossless medium simply from an extension of our understanding of wave propagation in free space. In general, for *lossless* isotropic media the permittivity is a *real* quantity which makes the wave vector a real quantity and the wave propagates without loss. The impedance is also a real quantity and the electric and magnetic fields are in phase.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 2.3:

An electromagnetic wave is propagating in a type of glass with an index of refraction $n_{\text{glass}} = 1.5$. The wave is propagating in a direction halfway between the positive x- and z-axes. If the free space wavelength is $\lambda_o = 628$ nm, find the wavelength in the glass and the wave vector. Also find an expression for the phasor form of the electric field in the glass if it has a maximum amplitude of 10 V/m oriented in the +y-direction.

Solution:

The wavelength in the glass is

$$\lambda_{\text{glass}} = \frac{\lambda_o}{n_{\text{glass}}} = \frac{628\,\text{nm}}{1.5} = 419\,\text{nm}$$

The propagation constant in the glass is

$$k = n_{\text{glass}} k_o = n_{\text{glass}} \frac{2\pi}{\lambda_o} = 1.5 \frac{2\pi}{628 \times 10^{-9}} = 1.5 \times 10^7 \quad \text{rad/m}$$

Referring to Figure 2.7, the unit vector in the direction of propagation is

$$\hat{k} = \frac{\hat{x} + \hat{z}}{\sqrt{2}}$$

The wave vector in the glass is then

$$\overline{k} = k\hat{k} = 1.5 \times 10^7 \frac{\hat{x} + \hat{z}}{\sqrt{2}} = 1.06 \times 10^7 \,(\hat{x} + \hat{z})$$

The complete phasor expression for the electric field is

$$\bar{E} = \bar{E}_o e^{-j\bar{k}\cdot r} = \hat{y}10e^{-j1.06\times 10^7(x+z)} \text{ V/m}$$

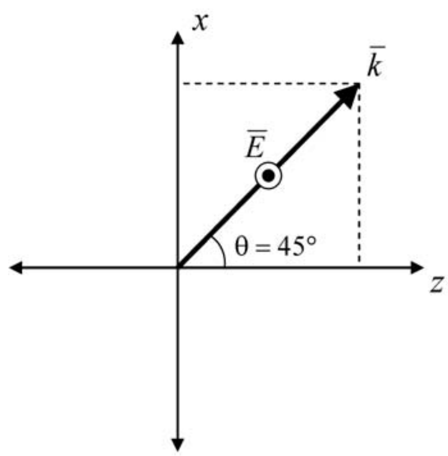


Figure 2.7: Direction of wave vector and electric field vector for Example 2.3.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

2.6.4 Plane Waves in Lossy Media

In this section, we investigate wave propagation in a medium that exhibits loss. Losses in a material are accounted for by defining a *complex* quantity to describe the permittivity. The complex permittivity causes the wave vector to also be a complex quantity, which in turn causes the plane wave to decay as it is propagating. In addition, the wave impedance becomes a complex quantity, which means that now the electric and magnetic fields are *out-of-phase*. We next investigate the mathematical description of propagation in lossy media.

One mechanism that leads to loss in a material is *conductive* loss, which occurs in materials that can conduct electric current, such as metals. If there are conductive losses in a material, then the electric field induces a current density given by

$$\bar{J} = \sigma\bar{E} \tag{2.73}$$

where σ is the conductivity in siemens per meter, or S/m (a siemen has the units of Ω^{-1}). Substituting this current density into Ampere's law (Equation 2.40) we obtain

$$\nabla \times \bar{H} = j\omega\varepsilon\bar{E} + \sigma\bar{E} = j\omega\varepsilon(1 - j\frac{\sigma}{\omega\varepsilon})\bar{E} \tag{2.74}$$

By defining a complex permittivity

$$\varepsilon_c \equiv \varepsilon\left(1 - j\frac{\sigma}{\omega\varepsilon}\right) \tag{2.75}$$

then we can rewrite Equation 2.74 as

$$\nabla \times \bar{H} = j\omega\varepsilon_c\bar{E} \tag{2.76}$$

Notice that this equation has exactly the same form as Ampere's law. In fact, *all the remaining equations hold by simply replacing the real quantity ε by the complex quantity ε_c.* This observation is extremely important! The complex wave impedance is

$$\eta_c = \sqrt{\frac{\mu_o}{\varepsilon_c}} = |\eta_c| e^{j\phi} \tag{2.77}$$

where the impedance has explicitly been written in polar form to show that now there will be an amplitude difference *and a phase difference* between the electric and magnetic fields. The complex propagation constant is

$$k_c = \omega\sqrt{\mu_o \varepsilon_c} \tag{2.78}$$

The complex wave vector is $\bar{k_c} = k_c\hat{k}$, where $\hat{k}$ is the unit vector in the direction of wave propagation. The time-harmonic Maxwell's equations in an infinite lossy medium are obtained by simply replacing $\bar{k}$ with $\bar{k_c}$:

$$\bar{k}_c \times \bar{E} = \omega\mu_o \bar{H} \tag{2.79}$$

$$\bar{k}_c \times \bar{H} = -\omega\varepsilon_c \bar{E} \tag{2.80}$$

$$\bar{k}_c \cdot \bar{E} = 0 \tag{2.81}$$

$$\bar{k}_c \cdot \bar{H} = 0 \tag{2.82}$$

which, with the exception of the complex wave vector, look identical to Equations 2.53-2.56. This means that the spatial relationships between the electromagnetic fields remain unchanged.

Using a similar derivation as in free space, it follows that for a homogeneous plane wave (one in which the wave is attenuating in the direction of propagation),

$$\bar{H} = \frac{1}{\eta_c}\hat{k} \times \bar{E} \tag{2.83}$$

It also follows that the Helmholtz equation in a lossy medium is

$$\nabla^2 \bar{E} + k_c^2 \bar{E} = 0 \tag{2.84}$$

One very important point to note when dealing with lossy materials is the following. For simplicity, let us assume that the direction of propagation is the z-direction. (Note that we could assume the propagation in any direction, but this assumption greatly simplifies the following analysis.) For propagation in the forward z-direction, the complex exponential can be written as

$$e^{-jk_c z} = e^{-j\omega\sqrt{\mu_o\varepsilon_c}z} = e^{-j(\beta \pm j\alpha)z} = e^{-j\beta z}e^{\pm\alpha z} \tag{2.85}$$

if we define $k_c = \omega\sqrt{\mu_o\varepsilon_c} = \beta \pm j\alpha$. Notice that there are two possible solutions since the square root of a complex number has two solutions. Which sign should we choose? Since the wave is propagating in the +z-direction, then the wave must decay in the +z-direction, and therefore we must choose $k_c = \beta - j\alpha$ so that $e^{-jk_c z} = e^{-j\beta z}e^{-\alpha z}$ and the wave exponentially decays in the +z-direction. If the wrong sign is chosen, the wave will exponentially grow in the +z direction, which does not make physical sense.

Representing ε_c as a vector in the complex plane, the tangent of the angle that the vector makes with the real axis is $\sigma/(\omega\varepsilon)$ (see Figure 2.8). This quantity is called the *loss tangent*:

$$\text{loss tangent} = \frac{\sigma}{\omega\varepsilon} \tag{2.86}$$

The loss tangent describes the loss characteristics of a medium. In a lossless material, $\sigma/(\omega\varepsilon) = 0$. If $\sigma/(\omega\varepsilon) \ll 1$, the material is called a *good dielectric*. If $\sigma/(\omega\varepsilon) \gg 1$, the material is called a *good conductor*,

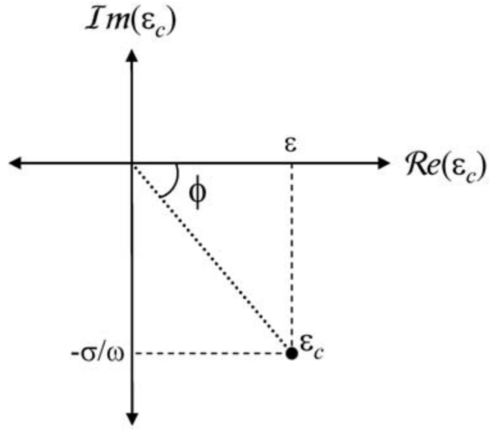


Figure 2.8: Plotting the complex permittivity, ε_c, as a vector in the complex plane, the loss tangent is the tangent of the angle that the vector makes with the real axis, i.e. $\tan\phi = \sigma/(\omega\varepsilon)$.

and if $\sigma/(\omega\varepsilon) = \infty$ the material is a *perfect conductor*. The concept of a "perfect conductor" is more of a mathematical idealization than an actual physical reality. Even so, highly conducting materials are often approximated as perfect conductors to simplify the mathematics while providing valuable physical insight into a problem.

Other definitions and terminology pertaining to propagation in lossy media that you will need to be familiar with are the following:

1. **Skin depth:** The skin depth, δ, is the distance in which the wave decays by e^{-1}. Since a wave propagating in a lossy medium decays as $e^{-\alpha z}$, we have $-1 = -\alpha\delta$, or $\delta = 1/\alpha$.

2. **Decibels:** Defined in terms of the electric field amplitude as $\text{dB} \equiv 20\log_{10}(E_2/E_1)$ or in terms of power as $\text{dB} \equiv 10\log_{10}(P_2/P_1)$. For example, if a certain wave has an initial amplitude of 10 V/m, but decays to 5 V/m after propagating 1 meter, then we have: $20\log_{10}(5/10) \approx -6$ dB/meter. Note that often the negative sign is dropped, and would simply be referred to as a "6 dB per meter loss."

3. **Nepers:** Defined as $\text{Nepers} \equiv \log_e(E_1/E_2)$. In a lossy medium, the loss in Nepers/meter$= \log_e(|E_1|/|E_1|e^{-\alpha}) = \log_e(e^{\alpha}) = \alpha$. Therefore, the attenuation constant α has the units Nepers/meter (Np/m).

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 2.4:

Human brain tissue has a relative permittivity and conductivity of $\varepsilon_r \approx 55$ and $\sigma \approx 2.1$ S/m, respectively, at 2.45 GHz [2]. This is a frequency of interest since many cell phones (and microwave ovens!) operate near this frequency. Assume that a plane wave at 2.45 GHz is normally incident from free space onto a human brain. Find the complex permittivity, complex impedance, and loss tangent.

Solution:

The complex permittivity of the brain tissue is

$$\varepsilon_{\text{brain}} = \varepsilon_r\varepsilon_o\left(1 - j\frac{\sigma}{\omega\varepsilon_r\varepsilon_o}\right)$$

$$= (55)\left(8.854\times10^{-12}\right)\left\{1 - j\frac{2.1}{(2\pi\times2.45\times10^9)(55)(8.854\times10^{-12})}\right\}$$

$$= (4.87 - j1.36) \times 10^{-10} \quad \text{F/m}$$

From this we obtain the complex impedance

$$\eta_{\text{brain}} = \sqrt{\frac{\mu_o}{\varepsilon_{\text{brain}}}} = \sqrt{\frac{4\pi \times 10^{-7}}{(4.87 - j1.36) \times 10^{-10}}}$$

$$= 49.4 + j6.77 \;\; \Omega = 49.9 e^{j0.136} \;\; \Omega$$

The loss tangent is

$$\text{loss tangent} = \frac{\sigma}{\omega \varepsilon_r \varepsilon_o} = \frac{2.1}{(2\pi \times 2.45 \times 10^9)\,(55)\,(8.854 \times 10^{-12})} = 0.280$$

$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

EXAMPLE 2.5:

Using the results from the previous example, find the phasor and time domain expressions for the electric and magnetic fields propagating in brain tissue at 2.45 GHz. Assume that the electric field vector points in the x-direction, propagates in the z-direction, and has a maximum amplitude of 10 V/m just below the surface of the brain tissue at $z = 0$. What is the skin depth at this frequency? After what depth has the field amplitude been attenuated by 100 dB?

Solution:

Because the brain tissue is a lossy material, the wave vector will be complex and the electric field will have the general form

$$\bar{E}(z) = E_o e^{-jk_{\text{brain}} z} \hat{x} = E_o e^{-j(\beta - j\alpha)z} \hat{x} = E_o e^{-j\beta z} e^{-\alpha z} \hat{x}$$

The radian frequency of the wave is $\omega = 2\pi \times 2.45 \times 10^9 = 1.54 \times 10^{10}$ rad/s. The complex propagation constant can then be found from the complex permittivity obtained in the previous example:

$$k_{\text{brain}} = \omega \sqrt{\mu_o \varepsilon_{\text{brain}}} = 1.54 \times 10^{10} \sqrt{(4\pi \times 10^{-7})\,(4.87 - j1.36) \times 10^{-10}}$$

$$= 384 - j52.7 = \beta - j\alpha$$

Using these values for β and α we can write the phasor and time-domain expressions for the electric field:

$$\bar{E}(z) = 10 e^{-j384z} e^{-52.7z} \hat{x} \;\; \text{V/m}$$

$$\bar{\mathcal{E}}(z,t) = 10 \cos(1.54 \times 10^{10} t - 384z) e^{-52.7z} \hat{x} \;\; \text{V/m}$$

The corresponding magnetic field phasor is found using the complex impedance found in the previous example:

$$\bar{H}(z) = \frac{1}{\eta_{\text{brain}}} \hat{z} \times \bar{E}(z) = \frac{\hat{z} \times \bar{E}(z)}{49.9 e^{j0.136}} = \frac{10 e^{-j384z} e^{-52.7z} \hat{y}}{49.9 e^{j0.136}}$$

$$= 0.201 e^{-j0.136} e^{-j384z} e^{-52.7z} \hat{y} \;\; \text{A/m}$$

Notice the phase shift (0.136 radians, or 7.79°) between the electric and magnetic fields. This phase shift is clearly apparent in the time domain form of the magnetic field:

$$\bar{\mathcal{H}}(z,t) = 0.201 \cos(1.54 \times 10^{10} t - 384z - 7.79^\circ) e^{-52.7z} \hat{y} \;\; \text{A/m}$$

The skin depth at 2.45 GHz is simply the inverse of the attenuation constant:

$$\delta = \frac{1}{\alpha} = \frac{1}{52.7} = 19.0 \ \text{ mm}$$

To find the depth at which the field has been attenuated by 100 dB, first note that a decrease in 100 dB corresponds to a field amplitude that has decreased to 1×10^{-5} of its initial value. Therefore, we find

$$1 \times 10^{-5} E_o = E_o e^{-\alpha z} \quad \rightarrow \quad z = -\ln(1 \times 10^{-5})/\alpha = 21.8 \,\text{cm}$$

Note that this requires a very large head!

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

2.6.5 Example of a Lossy Medium: Wave Propagation in Seawater

As an example of propagation in a lossy medium, let us look at a wave propagating in seawater at different frequencies. This example will show that one needs to be careful in making general conclusions about wave propagation in a lossy medium.

Seawater has a permittivity of $\varepsilon \approx 79\varepsilon_o$ and an electrical conductivity of $\sigma \approx 4$ S/m (siemens/meter; a siemen has the units of Ω^{-1}). The loss tangent for seawater is plotted as function of frequency in Figure 2.9. For frequencies less than 100 MHz (10^8 Hz), the loss tangent is greater than 10 and the seawater behaves like a good conductor. For frequencies greater than 10 GHz (10^{10} Hz), the loss tangent is less than 0.1 and seawater behaves like a good dielectric.

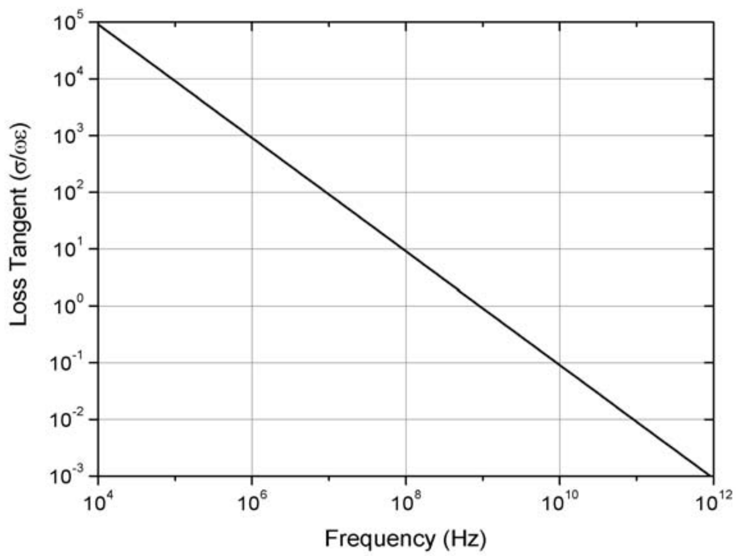


Figure 2.9: Loss tangent of seawater.

In Figure 2.10, the loss in decibels per meter (dB/m) is plotted for a wave in seawater as a function of frequency. Compare the results in Figure 2.9 and Figure 2.10 and note that even though seawater is a good conductor for $f < 100$ MHz and a good dielectric for $f > 10$ GHz, the loss per meter in the "good

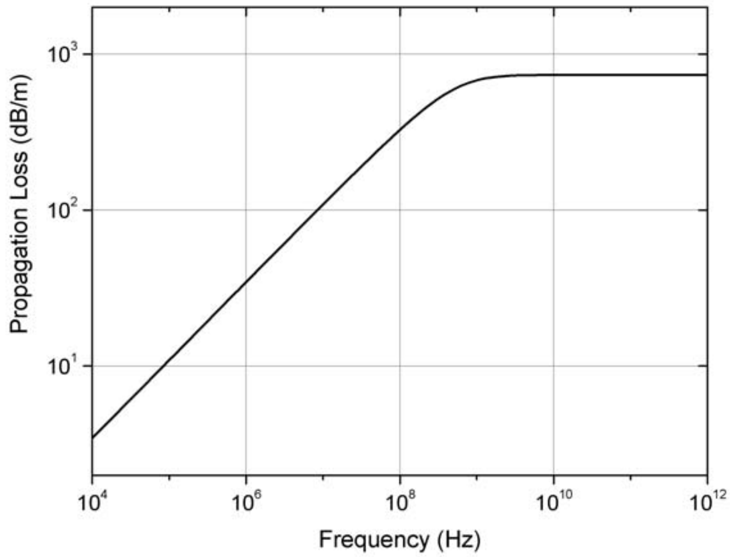


Figure 2.10: Propagation loss per meter in seawater.

conducting" region is lower than in the "good dielectric" region! Therefore, we cannot make the generalization that propagation loss is directly proportional to the conductivity of the medium.

We will next use the approximation of a good dielectric and good conductor to get a general feel for wave propagation in each type of medium. Note that the wave properties can always be analyzed using the exact complex permittivity, without making any approximations. However, we can develop insights by looking at these two limiting cases. They are:

1. **A good dielectric:** A material in which $\sigma/(\omega\varepsilon) \ll 1$
2. **A good conductor:** A material in which $\sigma/(\omega\varepsilon) \gg 1$

2.6.6 Plane Waves in a Good Dielectric

In a good dielectric, we use the fact that $\sigma/(\omega\varepsilon) \ll 1$ to make certain approximations. Using the power series approximation $(1-x)^{\frac{1}{2}} \approx 1 - x/2$ (which holds if $|x| \ll 1$), the approximate propagation constant k_c is

$$k_c = \omega\sqrt{\mu_o\varepsilon}\sqrt{1 - j\frac{\sigma}{\omega\varepsilon}} \approx \omega\sqrt{\mu_o\varepsilon}\left(1 - j\frac{\sigma}{2\omega\varepsilon}\right) = \omega\sqrt{\mu_o\varepsilon} - j\frac{\sigma}{2}\sqrt{\frac{\mu_o}{\varepsilon}} = \beta - j\alpha \tag{2.87}$$

Since $\sigma/(\omega\varepsilon) \ll 1$ it follows that α is very small in a good dielectric. In addition, the real portion of the propagation constant, β, is approximately equal to the value we would observe if the conductivity of the material were zero. The skin depth in a good dielectric is approximately

$$\delta = \alpha^{-1} \approx \frac{2}{\sigma}\sqrt{\frac{\varepsilon}{\mu_o}} \tag{2.88}$$

which can be quite large in a good dielectric. Because the loss is small, the wave penetrates a large distance into the material. The wave impedance in a good dielectric, η_c, can be approximated as

$$\eta_c = \sqrt{\frac{\mu_o}{\varepsilon}}\left(1 - j\frac{\sigma}{\omega\varepsilon}\right)^{-1/2} \approx \sqrt{\frac{\mu_o}{\varepsilon}}\left(1 + j\frac{\sigma}{2\omega\varepsilon}\right) \tag{2.89}$$

Again, since $\sigma/(\omega\varepsilon) \ll 1$, the imaginary portion of η_c is very small and the electric and magnetic fields are nearly in-phase.

2.6.7 Plane Waves in a Good Conductor

In a good conductor, $\sigma/(\omega\varepsilon) \gg 1$ and the propagation constant k_c is approximated as

$$k_c \approx \omega\sqrt{\mu_o\varepsilon}\sqrt{-j\frac{\sigma}{\omega\varepsilon}} = \sqrt{\omega\mu_o\sigma}\left(\frac{1-j}{\sqrt{2}}\right) = \sqrt{\frac{\omega\mu_o\sigma}{2}} - j\sqrt{\frac{\omega\mu_o\sigma}{2}} = \beta - j\alpha \tag{2.90}$$

(refer to Appendix B for details of this derivation). Notice that $\beta \approx \alpha$ in a good conductor. This is a very useful observation since, for example, we can obtain the conductivity by either knowing the propagation constant or the attenuation constant. Using this approximation, the skin depth in a good conductor is

$$\delta = \alpha^{-1} \approx \sqrt{\frac{2}{\omega\mu_o\sigma}} \tag{2.91}$$

Notice that the skin depth decreases as the frequency increases.

The wave impedance in a good conductor is approximated in a similar manner as

$$\eta_c = \sqrt{\frac{\mu_o}{\varepsilon_c}} = \sqrt{\frac{\mu_o}{\varepsilon[1 - j\sigma/(\omega\varepsilon)]}}$$

$$\approx \sqrt{\frac{\mu_o}{\varepsilon[-j\sigma/(\omega\varepsilon)]}} = \sqrt{\frac{\mu_o\omega}{\sigma}}\left(\frac{1+j}{\sqrt{2}}\right) = |\eta_c|e^{+j\phi} \tag{2.92}$$

Notice that the maximum phase angle of the impedance, ϕ, is 45°. This implies that the maximum phase shift between the electric and magnetic fields in a lossy medium is 45°.

Let us examine what happens if we let the conductivity approach infinity, i.e. if the medium becomes a *perfect conductor*. Allowing $\sigma \to \infty$ in Equation 2.92, we find that the *impedance approaches zero in a perfect conductor*. This means that a finite electric field will generate an infinite magnetic field (refer to Equation 2.83)! This obviously doesn't make sense. The only way that Equation 2.83 can still hold is if *all of the electromagnetic fields are zero inside a perfect conductor*. Another way to see this is from the skin depth (Equation 2.91). As $\sigma \to \infty$, the skin depth approaches zero. This means that any field that does enter a perfect conductor will instantly decay to zero, so that no electromagnetic fields can exist inside a perfect conductor! Note that even though a true "perfect conductor" does not exist in reality, in many cases very highly conducting materials such as metals can be approximated as perfect conductors at certain frequencies to simplify calculations and obtain reasonably accurate results.

$\ll\ll\ll\ll\ll\ll\ll\gg\gg\gg\gg\gg\gg\gg$

EXAMPLE 2.6:

At 300 MHz, copper is a good conductor with a conductivity of $\sigma = 5.8 \times 10^7$ S/m. At what depth in copper will an electric field have decayed to 10% of its initial amplitude? What is the wavelength of the electric field in the copper?

Solution:

For a good conductor

$$\alpha \approx \sqrt{\frac{\omega\mu_o\sigma}{2}} = \sqrt{\frac{(2\pi \times 300 \times 10^6)(4\pi \times 10^{-7})(5.8 \times 10^7)}{2}} = 2.62 \times 10^5 \quad \text{Np/m}$$

In a conductor, the amplitude of the electric field will decay as $E_o e^{-\alpha z}$, and we want to find the distance at which the amplitude has decayed to 10% of E_o:

$$0.1E_o = E_o e^{-\alpha z} \quad \to \quad e^{-\alpha z} = 0.1$$

Solving for z:

$$z = \frac{-\ln(0.1)}{\alpha} = \frac{2.30}{2.62 \times 10^5} = 8.79\mu m$$

Using the good conductor approximation we know that $\alpha \approx \beta$, and in the copper $\beta = 2\pi/\lambda$. Solving for λ:

$$\lambda = \frac{2\pi}{\beta} = \frac{2\pi}{2.62 \times 10^5} = 24.0\mu m$$

$\ll\ll\ll\ll\ll\ll\ll\ll\gg\gg\gg\gg\gg\gg\gg\gg$

2.6.8 Plane Waves in Dispersive Media

In electromagnetics, we are often interested in studying the effects of electromagnetic pulses propagating through devices. A pulse that has a finite width in the time domain can be decomposed into a spectrum of harmonics at different frequencies using a Fourier series. If these harmonics each propagate at a different velocity, then the pulse shape will change as it progresses through the device. This phenomenon is called *dispersion*. We will study this phenomenon in greater detail in Chapter 5. For now, we wish to investigate the properties of a *dispersive medium*, which is a material that causes electromagnetic waves at different frequencies to propagate at different velocities.

Most materials exhibit some degree of dispersive behavior. One example of a dispersive medium is a plasma. A plasma is a collection of oscillating positively and negatively charged particles in which the time-averaged charge density is zero. Examples of plasmas are found in the ionosphere and on the surface of the sun. In addition, highly conductive metals such as aluminum and gold behave like plasmas at optical frequencies (approximately $10^{14} \sim 10^{15}$ Hz).

We will now derive an approximation for the dielectric constant of a plasma. Let us assume that the mass of the positively charged particles (ions) is much larger than the negatively charged particles (electrons), there are no collisions between the particles, and the particles are not strongly bound to each other. As we did when considering a polarizable medium, we define a macroscopic polarization $\bar{\mathcal{P}}(t) \approx Nq\bar{x}(t)$, where N is the density of particles, q is the electron charge, and $\bar{x}(t)$ is the electron position, which is time-harmonic. We can then derive the force between the particles

$$\bar{\mathcal{F}}(t) = m_e \bar{a}(t) = m_e \frac{d^2 \bar{x}(t)}{dt^2} \tag{2.93}$$

where m_e is the electron mass. Since we are working with time-harmonic quantities, we can convert to the expressions to phasor form. Converting the position $\bar{x}(t)$ to a phasor $\bar{x}$, and the force $\bar{\mathcal{F}}(t)$ to a phasor $\bar{F}$, we can make the phasor transformation

$$\bar{\mathcal{F}}(t) = m_e \frac{d^2 \bar{x}(t)}{dt^2} \quad \Longrightarrow \quad \bar{F} = m_e (j\omega)^2 \bar{x} = -m_e \omega^2 \bar{x} \tag{2.94}$$

Assuming that there is no restoring force, there is only an inertia term and $\bar{F} = q\bar{E}$. Solving for $\bar{x}$ we obtain $\bar{x} = -q\bar{E}/m_e\omega^2$, which yields a macroscopic polarization (in phasor form) of

$$\bar{P} \approx -\frac{Nq^2 \bar{E}}{m_e \omega^2} \tag{2.95}$$

For a plasma, then, the constitutive relationship between the electric field and flux is

$$\bar{D} = \varepsilon_o \bar{E} + \bar{P} = \varepsilon_o \left(1 - \frac{Nq^2}{\varepsilon_o m_e \omega^2}\right) \bar{E} = \varepsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right) \bar{E} \tag{2.96}$$

where the plasma frequency is $\omega_p = q\sqrt{N/(\varepsilon_o m_e)}$. The plasma permittivity is then defined as

$$\varepsilon_p = \varepsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right) \tag{2.97}$$

such that $\bar{D} = \varepsilon_p \bar{E}$. Note that the permittivity is highly dependent upon frequency. The propagation constant of the plasma, k_p, is

$$k_p = \omega\sqrt{\mu_o \varepsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right)} = k_o\sqrt{1 - \frac{\omega_p^2}{\omega^2}} \tag{2.98}$$

The functional dependence of k_p on ω is shown in Figure 2.11. The dashed line corresponds to the propagation constant in free space, $k_o = \omega\sqrt{\mu_o \varepsilon_o}$. This result shows that as $\omega \to \infty$, the propagation constant approaches that of a wave in free space. Notice that only if the frequency is greater than the plasma frequency (if $\omega > \omega_p$) is the propagation constant a real quantity. If $\omega < \omega_p$, then $\varepsilon_p < 0$ (note that the permittivity can be negative in a material!) and the propagation constant is purely imaginary. Recall that an imaginary propagation constant corresponds to a wave that is exponentially decaying; therefore, an electromagnetic wave is attenuated in a plasma if the frequency is below the plasma frequency. As will be seen later, the propagation characteristics of electromagnetic fields in metallic waveguides are similar in some ways to wave propagation in plasmas.

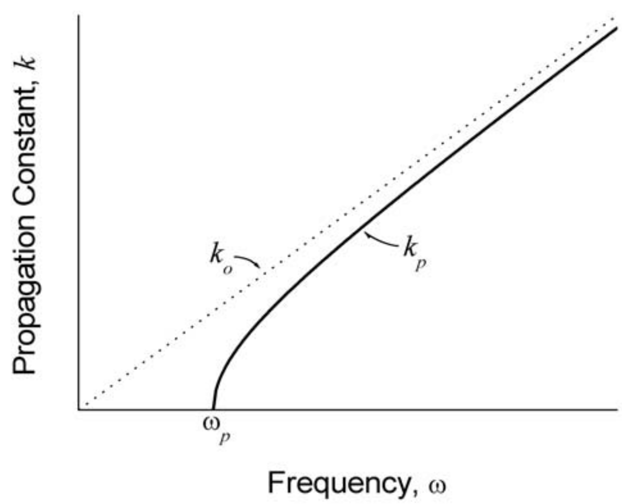


Figure 2.11: Propagation constant of an electromagnetic wave in a plasma (k_p). As $\omega \to \infty$, k_p approaches that of a wave in free space (k_o). The graph only shows the region in which the propagation constant is a real quantity ($\omega > \omega_p$). If $\omega < \omega_p$, then k_p is purely imaginary and the wave is attenuated.

2.7 Energy Flow and the Poynting Vector

Consider the total electromagnetic energy flowing through a closed region of space. If the difference between the power flowing in and power flowing out through the closed surface is non-zero, then the energy within the closed volume either decreases (the power is dissipated) because of loss, or the energy is increased in the region by storing energy in a magnetic or an electric field. How can we quantify this net power gain or loss in a region where electromagnetic fields are present?

In circuit theory, the power dissipated in a circuit element is found simply by multiplying the voltage across an element by the current flowing through it (volts × amps = watts). The question is, can we do a similar operation for the electromagnetic fields? Multiplying the units of the electric and magnetic field intensities we obtain (V/m)×(A/m) = W/m^2. From this standpoint we seem to be on the right track. Because both the electric and magnetic fields are vector quantities, it is reasonable to assume that the resulting power flow should also be a vector. The only way to multiply two vectors and obtain a vector is the cross-product: $\bar{\mathcal{E}} \times \bar{\mathcal{H}}$.

Recall that the physical significance of the divergence of a vector field, $\nabla \cdot \bar{A}$, is that it describes the rate at which the "field density" exits a given region of space. Since we are interested in the power flow in a region of space, we should consider calculating $\nabla \cdot (\bar{\mathcal{E}} \times \bar{\mathcal{H}})$. Applying the vector identity $\nabla \cdot (\bar{A} \times \bar{B}) = \bar{B} \cdot (\nabla \times \bar{A}) - \bar{A} \cdot (\nabla \times \bar{B})$, this can be written as

$$\nabla \cdot \left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) = \bar{\mathcal{H}} \cdot \left(\nabla \times \bar{\mathcal{E}}\right) - \bar{\mathcal{E}} \cdot \left(\nabla \times \bar{\mathcal{H}}\right) \tag{2.99}$$

Substituting Faraday's law and Ampere's law (Equations 2.2 and 2.3) we obtain

$$\nabla \cdot \left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) = \bar{\mathcal{H}} \cdot \left(-\frac{\partial \bar{\mathcal{B}}}{\partial t}\right) - \bar{\mathcal{E}} \cdot \left(\frac{\partial \bar{\mathcal{D}}}{\partial t} + \bar{\mathcal{J}}\right) \tag{2.100}$$

We are interested in the power flowing into a volume, so we integrate this expression over a given volume, V, to obtain

$$\int_V \nabla \cdot \left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) dV = \int_V \left(-\bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{B}}}{\partial t} - \bar{\mathcal{E}} \cdot \frac{\partial \bar{\mathcal{D}}}{\partial t}\right) dV - \int_V \left(\bar{\mathcal{E}} \cdot \bar{\mathcal{J}}\right) dV \tag{2.101}$$

We can now apply Gauss' theorem (also known as Green's theorem) which relates the volume integral of the divergence $\nabla \cdot \bar{A}$ to a surface integral of $\bar{A}$ over the boundary containing that volume (see Chapter 1):

$$\int_V \left(\nabla \cdot \bar{A}\right) dV = \oint_S \left(\bar{A} \cdot \hat{n}\right) dS \tag{2.102}$$

Applying Gauss' theorem and rearranging, we find

$$-\int_V \left(\bar{\mathcal{E}} \cdot \bar{\mathcal{J}}\right) dV = \int_V \left(\bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{B}}}{\partial t} + \bar{\mathcal{E}} \cdot \frac{\partial \bar{\mathcal{D}}}{\partial t}\right) dV + \oint_S \left[\left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) \cdot \hat{n}\right] dS \tag{2.103}$$

Assuming a linear, time-invariant medium, μ_o is not a function of time, and therefore

$$\bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{B}}}{\partial t} = \mu_o \left(\bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{H}}}{\partial t}\right) \tag{2.104}$$

Next, notice that, by using the chain rule, we can obtain the relationship

$$\frac{\partial}{\partial t}\left(\bar{\mathcal{H}} \cdot \bar{\mathcal{H}}\right) = \frac{\partial \bar{\mathcal{H}}}{\partial t} \cdot \bar{\mathcal{H}} + \bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{H}}}{\partial t} = 2\left(\bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{H}}}{\partial t}\right) \tag{2.105}$$

Therefore,

$$\mu_o \left(\bar{\mathcal{H}} \cdot \frac{\partial \bar{\mathcal{H}}}{\partial t}\right) = \frac{\mu_o}{2} \frac{\partial}{\partial t}\left(\bar{\mathcal{H}} \cdot \bar{\mathcal{H}}\right) \tag{2.106}$$

Furthermore, since the volume is not changing over time, the derivative with respect to time can be brought outside of the volume integral. Finally, applying the constitutive relationships and assuming that the volume does not change with time, we arrive at the *Poynting theorem* for linear, time-invariant, isotropic media:

$$-\int\limits_V \left(\bar{\mathcal{E}} \cdot \bar{\mathcal{J}}\right) dV = \frac{\partial}{\partial t} \int\limits_V \left(\frac{1}{2}\mu_o \bar{\mathcal{H}} \cdot \bar{\mathcal{H}} + \frac{1}{2}\varepsilon_o \bar{\mathcal{E}} \cdot \bar{\mathcal{E}}\right) dV + \oint\limits_S \left[\left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) \cdot \hat{n}\right] dS \tag{2.107}$$

Upon inspection of the Poynting theorem, the following observations are made:

1. The power either generated by the source $\bar{\mathcal{J}}$ or dissipated by the current $\bar{\mathcal{J}}$ in the volume V is $-\int\limits_V (\bar{\mathcal{E}} \cdot \bar{\mathcal{J}}) dV$.

2. The stored magnetic energy density in the volume V is $\frac{1}{2}\mu_o \bar{\mathcal{H}} \cdot \bar{\mathcal{H}}$.

3. The stored electric energy density in the volume V is $\frac{1}{2}\varepsilon_o \bar{\mathcal{E}} \cdot \bar{\mathcal{E}}$

4. The rate at which stored energy *increases* in the volume V is $\frac{\partial}{\partial t} \int\limits_V (\frac{1}{2}\mu_o \bar{\mathcal{H}} \cdot \bar{\mathcal{H}} + \frac{1}{2}\varepsilon_o \bar{\mathcal{E}} \cdot \bar{\mathcal{E}}) dV$.

It is therefore concluded that

$$\mathcal{P}(t) = \oint\limits_S \left[\left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) \cdot \hat{n}\right] dS \tag{2.108}$$

is the *instantaneous power flowing out of the volume* (note that by definition the unit vector $\hat{n}$ points outward, so we are looking at the power flowing outward in the direction of the unit vector). We then recognize

$$\bar{\mathcal{S}} = \bar{\mathcal{E}} \times \bar{\mathcal{H}} \tag{2.109}$$

as the *instantaneous power flow per unit area* (W/m^2). The vector quantity $\bar{\mathcal{S}}$ is called the *Poynting vector*.

Upon inspection, this definition may seem troubling. This is discussed in the following excerpt from *Fields and Waves in Communication Electronics* [3]: "Although it is known from the proof only that total energy flow out of a region per unit time is given by the total surface integral (Equation 2.108), it is often convenient to think of the vector $\bar{\mathcal{S}}$ defined by Equation 2.109 as the vector giving direction and magnitude of energy flow density at any point in space. Though this step does not follow strictly, it is a most useful interpretation and one which is justified for the majority of applications." For our purposes in this course, Equation 2.109 provides an accurate and useful definition of the power flow density at any instant in time.

More often, however, we are not interested in the power flow at a particular instant in time. Instead, the *average* power is often a more useful quantity. For example, imagine calculating the power required to thaw meat in a microwave oven. We are interested in how much power on average is deposited, and care less about the power at any instant. A more useful quantity is therefore the *time-averaged power flow per unit area* which is given by

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{1}{T} \int_0^T \left(\bar{\mathcal{E}} \times \bar{\mathcal{H}}\right) dt \tag{2.110}$$

and also has the units W/m^2. As it was shown in Chapter 1, this formidable time-average operation can be easily calculated if the phasor quantities are known:

$$\left\langle \bar{\mathcal{S}} \right\rangle = \left\langle \bar{\mathcal{E}} \times \bar{\mathcal{H}} \right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{E} \times \bar{H}^*\right\} = \frac{1}{2}\mathcal{R}e\left\{\bar{S}_c\right\} \tag{2.111}$$

where $\bar{S}_c = \bar{E} \times \bar{H}^*$ is the *complex Poynting vector*. In a typical analysis the phasor quantities are known. Instead of a integrating sines and cosines, the average power flow density is calculated by performing a cross-product using the phasors, then extracting the real part and dividing by 2. (See how simple this is when using phasors?)

For a plane wave of the form $\bar{E} = \bar{E}_o e^{-jkz}$ propagating in a *lossless* medium in the z-direction, it can be shown that the time-averaged power flow is (see Appendix B):

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{|\bar{E}_o|^2}{2\eta}\hat{z} \tag{2.112}$$

where

$$|\bar{E}_o| = \sqrt{\bar{E}_o \cdot \bar{E}_o^*} = \sqrt{|E_{\text{xo}}|^2 + |E_{\text{yo}}|^2 + |E_{\text{zo}}|^2} \tag{2.113}$$

since $\bar{E}_o$ may be, in general, a complex vector (refer to Chapter 1). Note that this relationship *only holds if the medium is lossless.* The time-averaged power flow for a plane wave in a *lossy* medium in the z-direction is (see Appendix B):

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{e^{-2\alpha z}|\bar{E}_o|^2}{2|\eta_c|}\cos(\phi)\hat{z} \tag{2.114}$$

where the complex impedance is written in polar form as $\eta_c = |\eta_c|e^{+j\phi}$. Compare this relationship to the lossless case. The loss in the material causes the power to exponentially decrease as the wave propagates. The complex permittivity also causes the electric and magnetic fields to be out-of-phase, which further reduces the power delivered by the electromagnetic wave. Electrical engineers that work on power distribution work fervently to maximize the factor $\cos(\phi)$ so that power is delivered efficiently. Considering what the angle ϕ represents, how can this be done?

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 2.7:

Assume that the following electromagnetic field is propagating in a vacuum:

$$\bar{E} = \hat{x}100e^{-j50z} \quad \text{V/m}$$

Find the average power density carried by this wave.

Solution:

Let's first work the problem using the Poynting vector, without using the simplified expression for lossless materials. The magnetic field associated with this electric field is

$$\bar{H} = \frac{1}{\eta_o}\hat{k} \times \bar{E} = \frac{\hat{z} \times \left(\hat{x}100e^{-j50z}\right)}{\eta_o} = \hat{y}\frac{100}{\eta_o}e^{-j50z}$$

The complex Poynting vector is

$$\bar{S}_c = \bar{E} \times \bar{H}^* = \left(\hat{x}100e^{-j50z}\right) \times \left(\hat{y}\frac{100}{\eta_o}e^{+j50z}\right) = \hat{z}\frac{100^2}{\eta_o}$$

The time-average power density is then

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{S}_c\right\} = \frac{1}{2}\mathcal{R}e\left\{\hat{z}\frac{100^2}{\eta_o}\right\} = \hat{z}\frac{100^2}{2\eta_o} = 13.3\hat{z} \quad \text{W/m}^2$$

Or, using the simplified expression for lossless media:

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{|\bar{E}_o|^2}{2\eta_o}\hat{z} = \frac{100^2}{2\eta_o}\hat{z} = 13.3\hat{z} \quad \text{W/m}^2$$

$\lll\lll\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 2.8:
An electromagnetic wave propagating in free space carries an average power of 5 mW/m^2. The electric field vector of the wave is given by

$$\bar{E} = E_i(4\hat{x} - j3\hat{y})e^{-j3000z} \quad \text{V/m}$$

What is the value of the constant E_i?

Solution:
The magnitude of the constant vector $\bar{E}_o$ is

$$|\bar{E}_o| = \sqrt{\bar{E}_o \cdot \bar{E}_o^*} = E_i\sqrt{(4\hat{x} - j3\hat{y}) \cdot (4\hat{x} + j3\hat{y})} = E_i\sqrt{16+9} = 5E_i$$

From Equation 2.112 we obtain

$$\left\langle \bar{S} \right\rangle = 0.005\hat{z} = \frac{(5E_i)^2}{2\eta_o}\hat{z} \quad \text{W/m}^2$$

From this we find $E_i = 0.02\sqrt{\eta_o} = 0.388$ V/m.

$\lll\lll\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg\ggg\ggg$

2.8 Polarization

The polarization of a wave is the geometric shape that the tip of the electric field vector traces at a fixed point in space as time progresses. In other words, it describes how the electric field vector changes with time at a fixed location. The polarization of a wave falls into one of three categories:

1. **Linear polarization:** The electric field vector oscillates only along one dimension. For example, if the electric field vector is confined to the x-axis then the wave is linearly polarized.

2. **Circular polarization:** The electric field "spirals" around as the wave propagates, tracing out a circle as time increases.

3. **Elliptical polarization:** Similar to circular polarization, except that the shape that electric field vector traces is an ellipse instead of a perfect circle. Both linear and circular polarization can be thought of as special cases of elliptical polarization: linear polarization occurs when the dimension of the ellipse along the minor axis is zero, and circular polarization occurs when the dimensions along the major and minor axes are equal.

Figure 2.12 illustrates the progression of the electric field vector for linear, circular, and elliptical polarization. Propagation in the positive z-direction is assumed in each case.

Determination of the polarization of a wave can often seem confusing. However, if you carry out the following systematic approach it is fairly straightforward:

1. To determine the polarization of a wave, you must look at the expression for the electric field of the wave *in the time domain.* Therefore, if the field is in phasor form you must convert to the time domain.

2. Pick a convenient *fixed location in space* at which to observe the wave. For example, $z = 0$ is often a convenient location for a wave propagating in the z-direction.

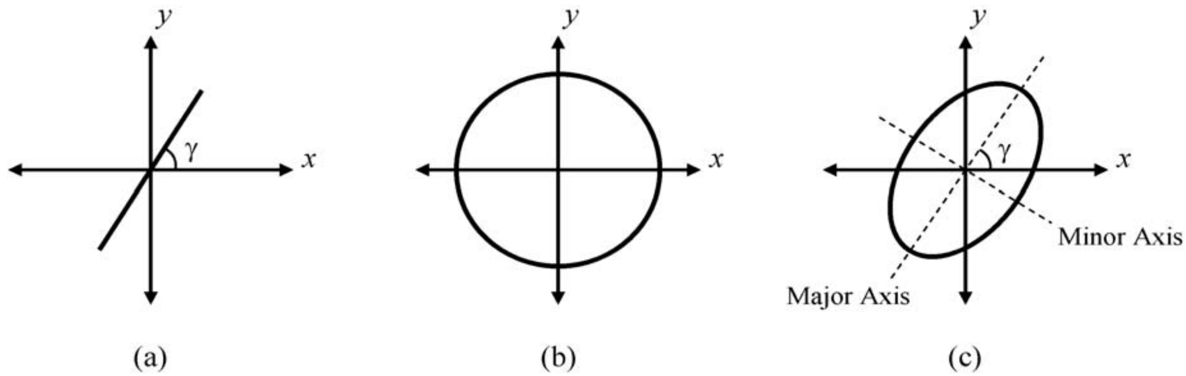


Figure 2.12: The polarization is determined by the shape that the electric field vector traces out in time as the wave propagates. (a) A linearly polarized wave, oriented at an angle γ with respect to the x-axis. (b) A circularly polarized wave. (c) An elliptically polarized wave, with the major axis oriented at an angle γ with respect to the x-axis. Linear and circular polarization can be thought of as special cases of elliptical polarization. Note that these figures assume wave propagation in the positive z-direction.

3. Determine the *magnitude and direction* of the electric field vector at some initial time, for example $t = 0$.

4. Increment time by some small value δt and determine how the electric field vector *progresses with time*. You may often choose $\omega t = \pi/2$, $\omega t = \pi$, $\omega t = 3\pi/2$, and so forth as convenient time increments at which to observe the electric field vector. Does the tip of the vector oscillate along a line (linear polarization) or rotate and trace out an ellipse or a circle?

5. If the electric field vector rotates and traces out an *ellipse* or a *circle*, then you must decide whether the wave is right-hand or left-hand polarized. To do this, point your right thumb in the direction of propagation and close your fingers. If your fingers are curled in the direction that the electric field vector rotates as time progresses, then the wave is *right-hand* elliptically or circularly polarized. This is illustrated in Figure 2.13. If the field vector is rotating in the *opposite* direction, then the wave is *left-hand* polarized.

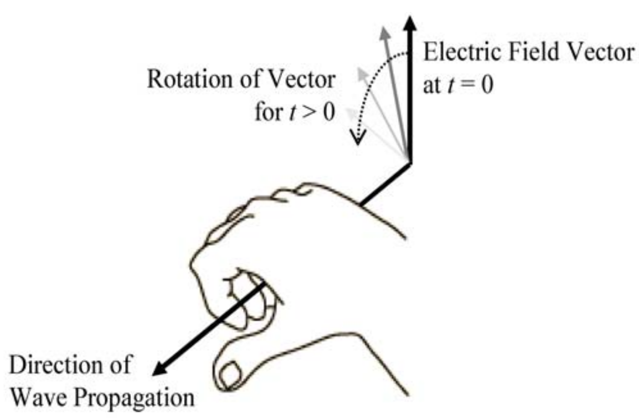


Figure 2.13: Determination of a right-hand polarized wave.

The ability to identify the polarization of a particular wave is very important in the study of electromagnetics. We next present specific examples of linearly, elliptically, and circularly polarized waves. In each of these examples, the wave propagation is in the positive z-direction. In general, however, this is not always the case, so you should be familiar with the process of identifying polarization for waves propagating in any direction.

2.8.1 Linearly Polarized Waves

Consider the following phasor electric field:

$$\bar{E}(z) = 2e^{-jkz}\hat{y} \tag{2.115}$$

Converting back to the time domain, the electric field is

$$\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{y} \tag{2.116}$$

This wave is travelling in the +z-direction (how do we know this?). We select the fixed location $z = 0$ at which to observe the wave. At $z = 0$, the electric field is

$$\bar{\mathcal{E}}(z=0,t) = 2\cos(\omega t)\hat{y} \tag{2.117}$$

At $t = 0$, the field vector is pointing along the +y-axis, and as time progresses it will only oscillate along this axis (see Figure 2.14). Therefore, this wave is *linearly polarized* along the y-axis.

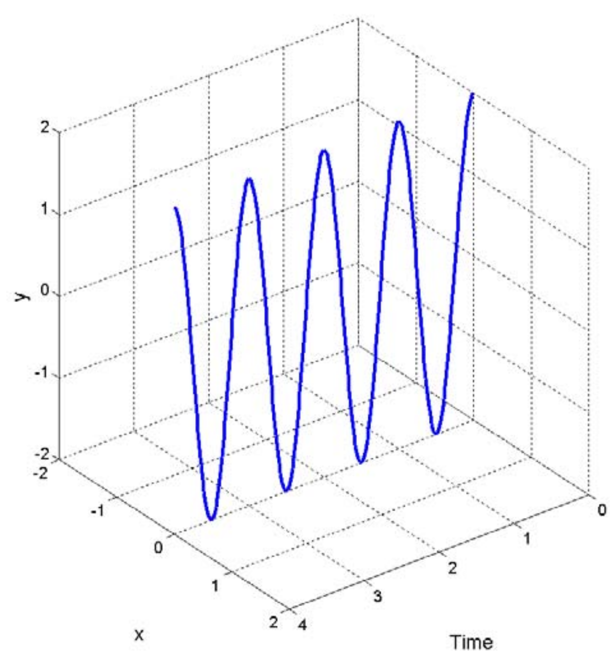


Figure 2.14: The progression of the electric field with time (fixed at $z = 0$) for the wave $\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{y}$. This wave is linearly polarized along the y-axis.

Does this mean that all linearly polarized waves must lie on the x- or y-axis? The answer is no. Consider the following wave, already in its time-domain form:

$$\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{x} - 2\cos(\omega t - kz)\hat{y} \tag{2.118}$$

At the observation point $z = 0$, the electric field is

$$\bar{\mathcal{E}}(z = 0, t) = 2\cos(\omega t)\hat{x} - 2\cos(\omega t)\hat{y} \tag{2.119}$$

At $t = 0$, the electric field vector is $\bar{\mathcal{E}} = 2\hat{x} - 2\hat{y}$. At $\omega t = \pi/2$, the electric field vector is $\bar{\mathcal{E}} = 0$. At $\omega t = \pi$, the electric field vector is $\bar{\mathcal{E}} = -2\hat{x} + 2\hat{y}$. Notice that since the x- and y-components of the electric field are in-phase, they trace out a line as shown in Figure 2.15. This wave is *linearly polarized*, but oriented at an angle of $\gamma = 135°$ with respect to the x-axis (refer to Figure 2.12).

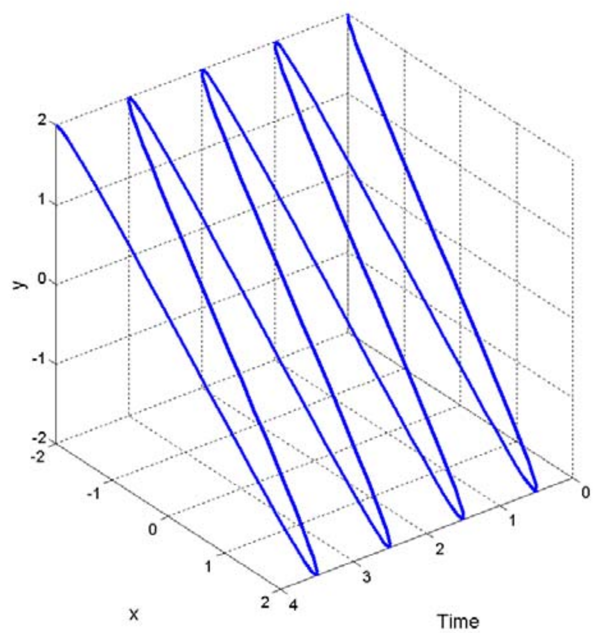


Figure 2.15: The progression of the electric field with time (fixed at $z = 0$) for the wave $\bar{\mathcal{E}}(z, t) = 2\cos(\omega t - kz)\hat{x} - 2\cos(\omega t - kz)\hat{y}$. The wave is oriented at an angle of $\gamma = 135°$ with respect to the x-axis.

2.8.2 Circularly Polarized Waves

Circular polarization only occurs when the two components of the electric field vector have equal magnitudes and are 90° out-of-phase with respect to each other. For example, consider the following electric field:

$$\bar{\mathcal{E}}(z, t) = 2\cos(\omega t - kz)\hat{x} + 2\sin(\omega t - kz)\hat{y} \tag{2.120}$$

Fixing our observation point at $z = 0$, we see that the field vector at $t = 0$ is $\bar{\mathcal{E}} = 2\hat{x}$. When we increment time to $\omega t = \pi/2$, the electric field vector is $\bar{\mathcal{E}} = 2\hat{y}$. At $\omega t = \pi$, the electric field vector is $\bar{\mathcal{E}} = -2\hat{x}$. If we continue this process, we see that the vector rotates as shown in Figure 2.16. Because the magnitude of the x- and y-components are equal, this describes circular polarization. Placing the thumb of your right hand in the direction of propagation (positive z-direction) and curling your fingers, you will find that the vector rotates in the direction of your curling fingers. Therefore, this wave is *right-hand circularly polarized*.

Now consider the electric field

$$\bar{\mathcal{E}}(z, t) = 2\cos(\omega t - kz)\hat{x} - 2\sin(\omega t - kz)\hat{y} \tag{2.121}$$

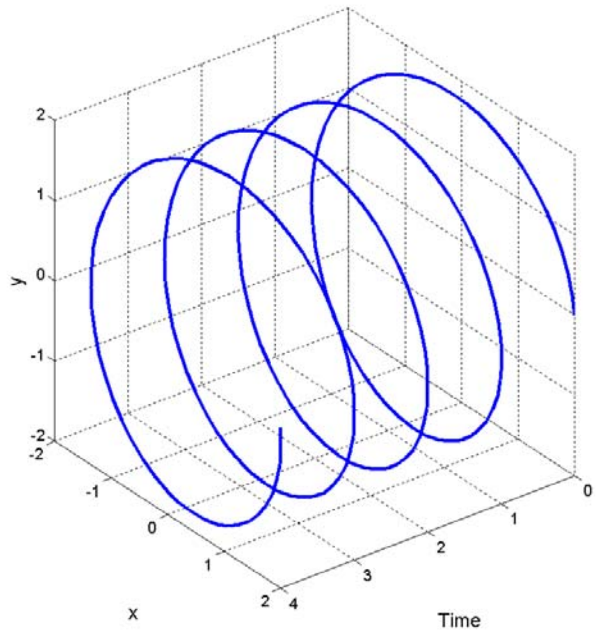


Figure 2.16: The progression of the electric field with time (fixed at $z = 0$) for the wave $\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{x} + 2\sin(\omega t - kz)\hat{y}$. This wave is right-hand circularly polarized.

When we fix our observation point at $z = 0$, we see that the field vector rotates in the opposite direction, as shown in Figure 2.17. In this case, you will find that when you point your left thumb in the positive z-direction, your fingers will curl in the direction of rotation. Therefore, this wave is *left-hand circularly polarized*.

2.8.3 Elliptically Polarized Waves

Since linear and circular polarization can be thought of as special cases of elliptical polarization, we obtain elliptical polarization any time that the wave *does not* exhibit linear or circular polarization. In other words, assuming propagation in the z-direction, elliptical polarization is obtained when the electric field has components in the x- and y-direction, but those components are not 90° out-of-phase with respect to each other and/or have different magnitudes. As an example, considering the following field:

$$\bar{\mathcal{E}}(z,t) = \cos(\omega t - kz)\hat{x} + 2\sin(\omega t - kz)\hat{y} \tag{2.122}$$

As before, we observe the wave at the fixed location $z = 0$:

$$\bar{\mathcal{E}}(z = 0,t) = \cos(\omega t)\hat{x} + 2\sin(\omega t)\hat{y} \tag{2.123}$$

At $t = 0$, the electric field vector is $\bar{\mathcal{E}} = \hat{x}$. When we increment time to $\omega t = \pi/2$, the electric field vector is $\bar{\mathcal{E}} = 2\hat{y}$. At $\omega t = \pi$, the electric field vector is $\bar{\mathcal{E}} = -\hat{x}$. Continuing the process, we find that the electric field vector traces out an ellipse, as shown in Figure 2.18. Placing your right thumb in the direction of propagation (+z), your fingers curl in the same direction as the rotation of the field with increasing time. Therefore, this wave is *right-hand elliptically polarized*, with the major axis of the ellipse aligned with the y-axis.

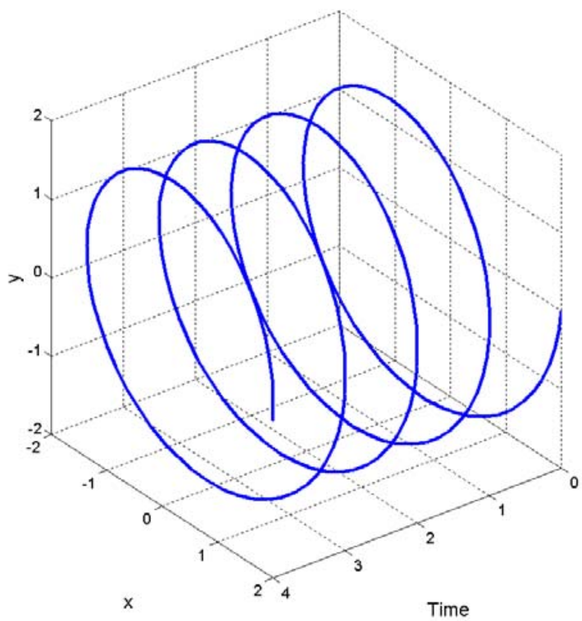


Figure 2.17: The progression of the electric field with time (fixed at $z = 0$) for the wave $\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{x} - 2\sin(\omega t - kz)\hat{y}$. This wave is left-hand circularly polarized.

Consider next the following field:

$$\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{x} - 2\cos(\omega t - kz - \pi/4)\hat{y} \tag{2.124}$$

At the fixed location $z = 0$ the field is

$$\bar{\mathcal{E}}(z = 0,t) = 2\cos(\omega t)\hat{x} - 2\cos(\omega t - \pi/4)\hat{y} \tag{2.125}$$

Now we see that at $t = 0$, the electric field vector is $\bar{\mathcal{E}} = 2\hat{x} - \sqrt{2}\hat{y}$. When we increment time to $\omega t = \pi/2$, the electric field vector is $\bar{\mathcal{E}} = -\sqrt{2}\hat{y}$. At $\omega t = \pi$, the electric field vector is $\bar{\mathcal{E}} = -2\hat{x} + \sqrt{2}\hat{y}$. Continuing the process, we find that the electric field vector traces out an ellipse as shown in Figure 2.19. The major axis of the ellipse is oriented at an angle of $\gamma = 135°$ with respect to the x-axis (refer to Figure 2.12). Checking the direction of rotation, we see that this wave is *left-hand elliptically polarized.*

As can be seen from the examples, the difficulty of specifying polarization increases as the fields become more complex. Determining the polarization of a wave is further complicated by the fact that the waves do not always propagate along the z-direction, and may not be given in terms of the electric field. In general, though, as long as the steps outlined previously are performed, the polarization will be accurately identified.

2.9 Important Concepts

Upon the completion of this chapter, you should be able to:

- Recall Maxwell's equations based on physical arguments.
- Recall the names of all the electromagnetic fields quantities and their units.

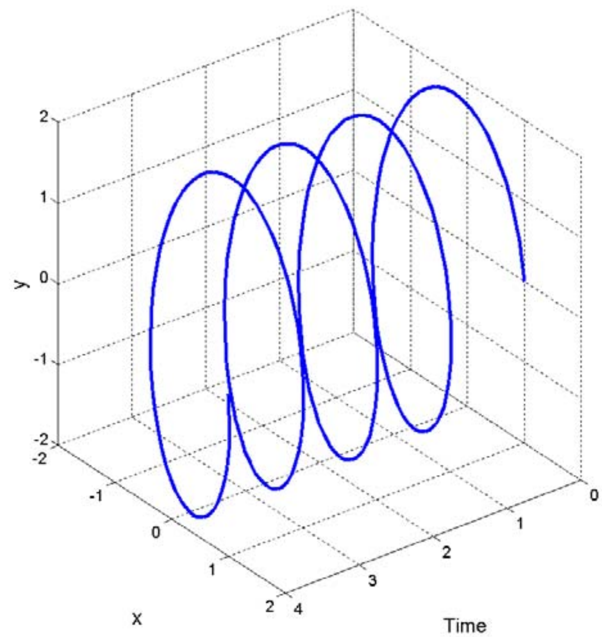


Figure 2.18: The progression of the electric field with time (fixed at $z = 0$) for the wave $\overline{\mathcal{E}}(z,t) = \cos(\omega t - kz)\hat{x} + 2\sin(\omega t - kz)\hat{y}$. This wave is right-hand elliptically polarized, with the major axis oriented along the y-axis.

- Derive the wave equations for the electric and magnetic fields based on recollection of Maxwell's equations and describe the underlying assumptions.
- Describe the concept of a plane wave and why the plane wave is fundamental to the general understanding of the propagation of electromagnetic waves.
- Understand what is meant by the wave vector, the "k-vector surface," and its physical interpretation.
- Explain what is meant by the following in relation to material properties: homogeneous/inhomogeneous, isotropic/anisotropic, linear/nonlinear, time variant/invariant, and dispersive/nondispersive.
- Understand the importance of the propagation constant and the wave impedance to describe wave propagation, regardless of the type of medium in which the wave is propagating.
- Recall the spatial and temporal relationships of the electric field intensity, magnetic field intensity and the wave vector in lossless media and lossy media.
- Derive the complex permittivity from first principles, and understand the motivation for defining a complex permittivity in a lossy medium.
- Understand the application of a complex wave vector and complex impedance to describe plane wave propagation in lossy medium, and how these quantities simplify to describe wave propagation in good dielectrics and good conductors. You should also be able to show how a good conductor approaches a perfect conductor in the limit as the conductivity approaches infinity.
- Define and apply the following quantities: loss tangent, skin depth, decibels and Nepers.

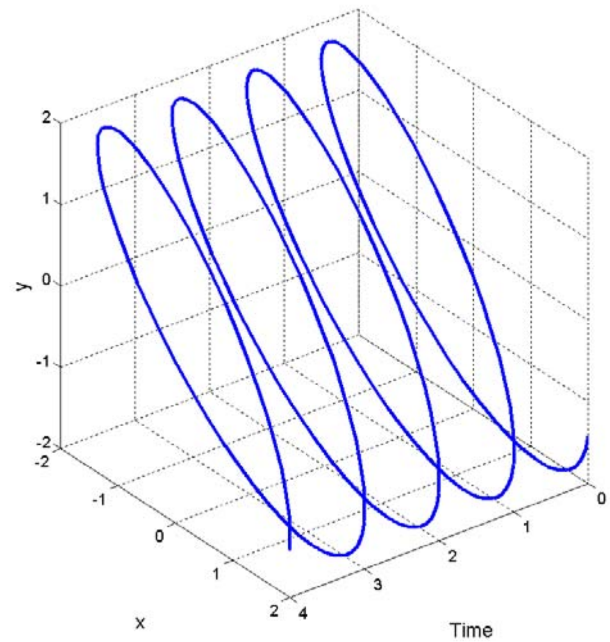


Figure 2.19: The progression of the electric field with time (fixed at $z = 0$) for the wave $\bar{\mathcal{E}}(z,t) = 2\cos(\omega t - kz)\hat{x} - 2\cos(\omega t - kz - \pi/4)\hat{y}$. The major axis of the ellipse is oriented at an angle of $\gamma = 135°$ with respect to the x-axis.

- Understand the Poynting vector and use the complex Poynting vector to determine the average power flow of a plane wave in lossless and lossy media.
- Define what is meant by the polarization of a wave, and be able to specify the polarization in terms of linear, circular and elliptical polarization.

Chapter 3

Reflection and Transmission of Waves

Heinrich Hertz (1857-1894)

"**March 19, 1888**: I would like to tell you of the progress as follows: in air, electrodynamic waves are reflected by solid conducting walls, reflected waves interfere with the incident waves - the incidence is vertical - and give rise to standing waves in the air. In the first wave length, calculated from the wall, these phenomena are very pronounced and varied, and I believe that we cannot reveal the undulatory nature of sound in free space as clearly as the undulatory nature of this electrodynamic propagation. The first node can be determined very accurately and we can measure the wavelength in an extremely direct manner. **November 30, 1888**: When you asked me in Berlin if I had carried other experiments on waves. I had nothing important to tell you at the time. Now, however, I have made some progress which firmly establishes the relationship between light and electricity and which I am anxious to tell you about."

—Extracts from letters by Heinrich Hertz to Hermann von Helmholtz

In the previous chapter, we focused on electromagnetic waves propagating in "infinite" media. The next step is to explore the consequences of a wave when it reaches a boundary between two dissimilar media. While this appears a simple problem, the behavior of electromagnetic waves at a boundary, or interface, is complex and full of surprises. This is a very important chapter, since most electromagnetic problems of interest to engineers involve waves propagating through or along interfaces. We begin the chapter discussing the boundary conditions for a wave at a single interface. This is then extended to a discussion of waves propagating through multiple interfaces.

3.1 Boundary Conditions

The boundary conditions relate the electromagnetic fields and fluxes on either side of an interface between two media. Let us refer to the two dissimilar material regions as "region 1" and "region 2." Assume that a wave propagating in region 1 is incident onto a boundary between regions 1 and 2. We first define a unit vector, $\hat{n}_{21}$, that is normal to the planar interface and points from region 2 into region 1, as shown in Figure 3.1. Although we will not show the details here, it can be shown that the electric and magnetic fields on either side of the interface must satisfy certain boundary conditions. (For a detailed derivation of the boundary conditions, refer to: D. K. Cheng. *Field and Wave Electromagnetics.* Addison-Wesley, 1989, Third Edition. pp. 116-121 & 329-330.)

The boundary conditions state that the *tangential* components of the electric fields (the components parallel to the interface) are continuous across the interface. In phasor form this condition is

$$\hat{n}_{21} \times \bar{E}_1 = \hat{n}_{21} \times \bar{E}_2 \tag{3.1}$$

The *tangential* components of the magnetic fields are related by

$$\hat{n}_{21} \times \bar{H}_1 = \hat{n}_{21} \times \bar{H}_2 + \bar{J}_s \tag{3.2}$$

where $\bar{J}_s$ is a surface current, with units of A/m. This equation states that the tangential components of the magnetic field differ across the interface by the amount $\bar{J}_s$. However, $\bar{J}_s = 0$ in all cases *except* when region 2 is a perfect conductor! If region 2 is perfectly conducting, the magnetic field in region 2 is zero ($\bar{H}_2 = 0$) which forces a surface current to exist.

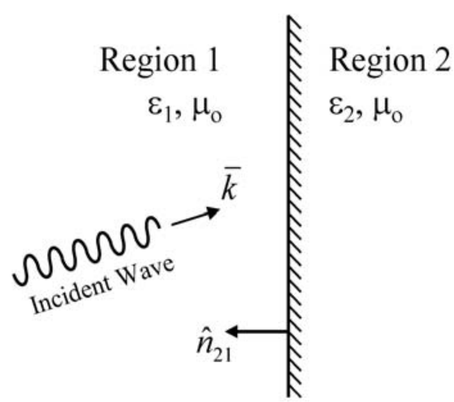


Figure 3.1: An electromagnetic wave incident on an interface between two media. The unit vector $\hat{n}_{21}$ is defined as pointing into region 1 (the incident region).

There are an additional two boundary conditions that relate the electromagnetic flux densities on either side of the interface. It can be shown that the *normal* components of the electric flux densities (the components perpendicular to the interface) are related by

$$\hat{n}_{21} \cdot \bar{D}_1 = \hat{n}_{21} \cdot \bar{D}_2 + \rho_s \quad (3.3)$$

where ρ_s is a surface charge density with units C/m^2. If the two regions are dielectrics, then $\rho_s = 0$ unless free charges exist at the interface between the two media. It can also be shown that the *normal* magnetic flux densities are continuous, i.e.

$$\hat{n}_{21} \cdot \bar{B}_1 = \hat{n}_{21} \cdot \bar{B}_2 \quad (3.4)$$

These four boundary conditions are more than we need to completely specify the fields at the interface. We can solve the interface problem using either the tangential or the normal boundary conditions. Because, in general, we do not know ρ_s it is easier to use the tangential conditions. For this reason, general solutions to interface problems are found by matching (setting equal) the *tangential electric fields* and the *tangential magnetic fields* at the interface between the two media.

If region 2 is a perfect conductor, then all of the electromagnetic fields in region 2 must be zero. This means that the boundary conditions at a *perfectly conducting interface* are:

$$\hat{n}_{21} \times \bar{E}_1 = 0 \quad (3.5)$$

$$\hat{n}_{21} \times \bar{H}_1 = \bar{J}_s \quad (3.6)$$

$$\hat{n}_{21} \cdot \bar{E}_1 = \rho_s/\varepsilon_1 \quad (3.7)$$

$$\hat{n}_{21} \cdot \bar{H}_1 = 0 \quad (3.8)$$

Notice that while the *tangential electric field* and the *normal magnetic field* must be zero at a perfectly conducting boundary, the *normal electric field* and *tangential magnetic field* are, in general, nonzero.

3.2 Normal Incidence on a Perfect Conductor

Assume that a plane wave propagating in free space in the z-direction is incident on a perfect conductor. For simplicity, we will focus on the case of normal incidence. We will align our coordinate system such that the interface between the free space region and the perfect conductor is the $z = 0$ plane. Without loss of generality, we can rotate our coordinate axes such that the electric field vector points in the y-direction and the incident wave propagates in the positive z-direction, as shown in Figure 3.2. Using phasor notation, we can write general expressions for the incident and reflected fields. In phasor form, the incident electric and magnetic fields are (recall that $\bar{H} = \hat{k} \times \bar{E}/\eta$, where η is the wave impedance of the medium)

$$\bar{E}_i = E_i e^{-jk_o z} \hat{y} \quad (3.9)$$

$$\bar{H}_i = \frac{-E_i}{\eta_o} e^{-jk_o z} \hat{x} = H_i e^{-jk_o z} \hat{x} \quad (3.10)$$

where E_i is the amplitude of the incident electric field and H_i is the amplitude of the incident magnetic field. The reflected fields propagate in the *negative* z-direction, and are therefore

$$\bar{E}_r = E_r e^{+jk_o z} \hat{y} \quad (3.11)$$

$$\bar{H}_r = \frac{E_r}{\eta_o} e^{+jk_o z} \hat{x} = H_r e^{+jk_o z} \hat{x} \quad (3.12)$$

where E_r and H_r are the amplitudes of the reflected electric and magnetic fields. Note that both the incident and reflected fields propagate through the free space region with the propagation constant $k_o = \omega\sqrt{\mu_o \varepsilon_o} = \omega/c$. Also, recall that the impedance of free space is $\eta_o = \sqrt{\mu_o/\varepsilon_o} \approx 377\ \ \Omega$.

Recall that we previously established that *no electric or magnetic fields can exist inside a perfect conductor*. Therefore the transmitted fields are

$$\bar{E}_t = 0 \tag{3.13}$$

$$\bar{H}_t = 0 \tag{3.14}$$

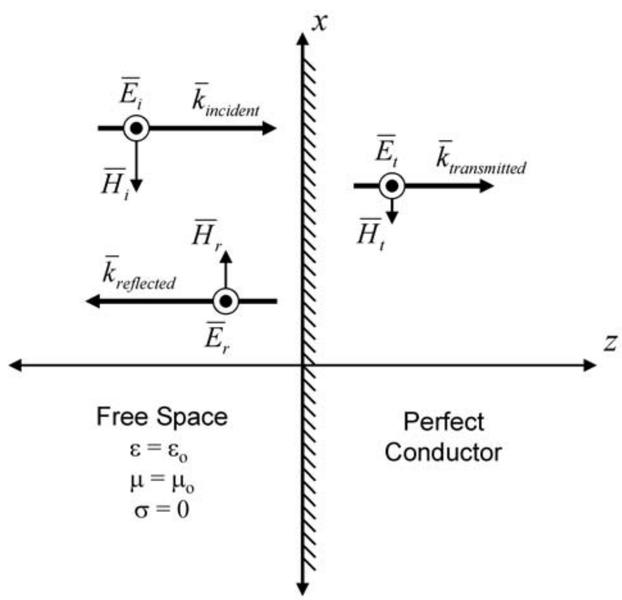


Figure 3.2: Orientation of the incident, reflected, and transmitted fields at the interface between free space and a perfectly conducting region, at normal incidence.

We must next relate the amplitude of the reflected field to the incident field amplitude. To do this, we apply the boundary condition for the tangential electric fields (Equation 3.5) at the interface between the two regions at $z = 0$:

$$-\hat{z} \times \bar{E}_i(z = 0) - \hat{z} \times \bar{E}_r(z = 0) = 0 \tag{3.15}$$

Substituting the expressions for the electric fields into this equation yields:

$$E_i + E_r = 0 \tag{3.16}$$

From this, we obtain an expression for the *electric field reflection coefficient* from a perfectly conducting interface at *normal incidence*:

$$\Gamma = \frac{E_r}{E_i} = -1 \tag{3.17}$$

Physically, this means that the electric field experiences a 180° phase shift upon reflection from a perfectly conducting interface, but no change in amplitude. The total electric field in region 1 is

$$\bar{E}_1 = E_i e^{-jk_o z}\hat{y} - E_i e^{+jk_o z}\hat{y} = E_i(e^{-jk_o z} - e^{+jk_o z})\hat{y}$$

$$= -j2E_i \sin(k_o z)\, \hat{y} \tag{3.18}$$

Transforming back to the time domain, the total electric field in region 1 is

$$\bar{\mathcal{E}}_1(z,t) = \mathcal{R}e\left\{-j2E_i \sin(k_o z)\, e^{j\omega t}\right\} \hat{y} = 2E_i \sin(k_o z) \sin(\omega t)\, \hat{y} \tag{3.19}$$

Notice that the total tangential electric field amplitude is always zero at $z = 0$, as expected. In addition, the total electric field is always zero if $k_o z$ is *any* multiple of π. These locations where the field is always zero do not change with time, but instead are fixed. This fixed portion is called a "standing wave," which we will discuss in greater detail later in the chapter.

What happens to the magnetic field when it is reflected from the perfectly conducting boundary? From Equations 3.10 and 3.12 we find that $H_i = -E_i/\eta_o$ and $H_r = E_r/\eta_o$. Using these results, we find that the magnetic field experiences a reflection

$$\frac{H_r}{H_i} = \frac{E_r/\eta_o}{-E_i/\eta_o} = -\Gamma = +1 \tag{3.20}$$

This means that the magnetic field experiences no phase shift or amplitude change upon reflection from a perfectly conducting interface at normal incidence! This means that the total magnetic field in region 1 is

$$\bar{H}_1 = H_i e^{-jk_o z}\hat{x} + H_i e^{+jk_o z}\hat{x} = H_i(e^{-jk_o z} + e^{+jk_o z})\hat{x}$$

$$= 2H_i \cos(k_o z)\, \hat{x} = \frac{-2E_i}{\eta_o} \cos(k_o z)\, \hat{x} \tag{3.21}$$

Transforming to the time-domain, the total magnetic field is

$$\bar{\mathcal{H}}_1(z,t) = \mathcal{R}e\left\{\frac{-2E_i}{\eta_o} \cos(k_o z)\, e^{j\omega t}\right\} \hat{x} = \frac{-2E_i}{\eta_o} \cos(k_o z) \cos(\omega t)\, \hat{x} \tag{3.22}$$

From these results, we make several observations:

1. The incident and reflected electric fields differ only in sign and direction of propagation.
2. The incident and reflected magnetic fields differ only in direction of propagation.
3. The total magnetic field at the interface of the perfect conductor ($z = 0$) is **not** zero! Instead, it has a magnitude $|\bar{H}_1(z = 0)| = 2E_i/\eta_o$. This means that there must be a surface current, $\bar{J}_s$, flowing at the interface (see Equation 3.6).

Let us now solve for the surface current flowing at the interface. From Equation 3.6 we find

$$\bar{J}_s = \hat{n}_{21} \times \bar{H}_1(z = 0) = -\hat{z} \times \left(\frac{-2E_i}{\eta_o}\hat{x}\right) = \frac{2E_i}{\eta_o}\hat{y} \tag{3.23}$$

This current exists only at the surface of the perfect conductor. Notice that the surface current at the interface always flows in the direction of the electric field. In this example, if thin slots are cut in the conducting plane in the y-direction the current will not be disturbed significantly. However, if a thin slot is cut in the x-direction, the current flow will be significantly disturbed, and hence the reflection properties of the surface will be altered greatly.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.1:

A 20 MHz plane wave in free space is normally incident on a perfectly conducting region. The interface between the free space and perfectly conducting regions is the $z = 0$ plane. The electric field has an amplitude of 100 V/m and is polarized in the x-direction. Find the phasor expressions for the incident and reflected electric and magnetic fields, and the time-domain expression for the total electric field in the free space region.

Solution:

The wave vector of the incident wave is

$$\bar{k}_o = \frac{\omega}{c}\hat{z} = \frac{2\pi \times 20 \times 10^6}{3 \times 10^8}\hat{z} = 0.4\hat{z}$$

Referring to Figure 3.3, the incident electric and magnetic fields are

$$\bar{E}_i = E_i e^{-j\bar{k}_o \cdot \bar{r}} = 100 e^{-j0.4z}\hat{x} \;\; \text{V/m}$$

$$\bar{H}_i = \frac{1}{\eta_o}\hat{k} \times \bar{E}_i = \left(\frac{1}{120\pi}\hat{z}\right) \times \left(100 e^{-j0.4z}\hat{x}\right) = \frac{5}{6\pi}e^{-j0.4z}\hat{y} \;\; \text{A/m}$$

For a perfectly conducting interface the electric field reflection coefficient is $\Gamma = -1$. Also, the reflected fields propagate in the -z direction, so the sign of the exponent must be changed. The reflected fields are

$$\bar{E}_r = \Gamma E_i e^{+j0.4z}\hat{x} = -100 e^{+j0.4z}\hat{x} \;\; \text{V/m}$$

$$\bar{H}_r = \frac{1}{\eta_o}\hat{k} \times \bar{E}_r = \frac{1}{120\pi}(-\hat{z}) \times \left(-100 e^{+j0.4z}\hat{x}\right) = \frac{5}{6\pi}e^{+j0.4z}\hat{y} \;\; \text{A/m}$$

Note that the reflected magnetic field could be found directly from the incident magnetic field using the magnetic field reflection coefficient of +1 at a perfect conductor. The total electric field in the free space region (region 1) is found by adding the incident and reflected fields:

$$\bar{E}_1 = \bar{E}_i + \bar{E}_r = 100 e^{-j0.4z}\hat{x} - 100 e^{+j0.4z}\hat{x} = 100\left(e^{-j0.4z} - e^{+j0.4z}\right)\hat{x}$$

$$= 100[\cos(0.4z) - j\sin(0.4z) - \cos(0.4z) - j\sin(0.4z)]\hat{x}$$

$$= -j200\sin(0.4z)\hat{x} \;\; \text{V/m}$$

Converting to the time-domain, where $\omega = 2\pi \times 20 \times 10^6 = 120 \times 10^6\pi$:

$$\bar{\mathcal{E}}_1(z,t) = \mathcal{R}e\left\{\bar{E}_1 e^{j\omega t}\right\} = \mathcal{R}e\left\{-j200\sin(0.4z)e^{j\omega t}\right\}\hat{x}$$

$$= \mathcal{R}e\left\{-j200\sin(0.4z)[\cos(\omega t) + j\sin(\omega t)]\right\}\hat{x}$$

$$= 200\sin(0.4z)\sin(120 \times 10^6\pi t)\hat{x} \;\; \text{V/m}$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.2:

Find the surface current, $\bar{J}_s$, flowing at the interface in the previous example.

Solution:

To find the surface current, we must first find the total magnetic field in region 1 at the boundary. The total magnetic field at the boundary is simply the addition of the incident and reflected fields at $z = 0$:

$$\bar{H}_1(z = 0) = \bar{H}_i(z = 0) + \bar{H}_r(z = 0)$$

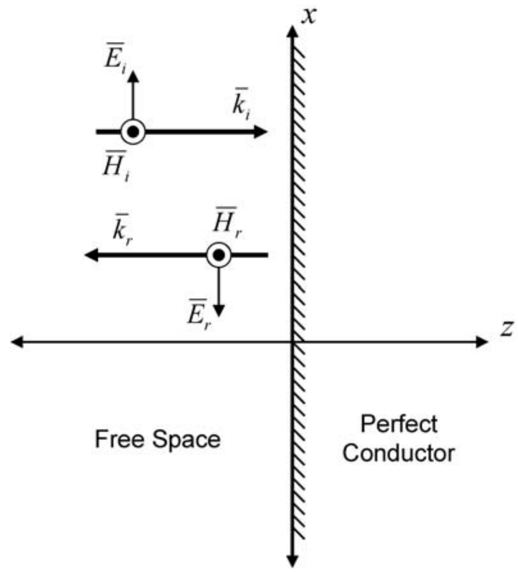


Figure 3.3: Incident and reflected field vectors for Example 1. Notice that there are no fields inside of the perfect conductor.

$$= \frac{5}{6\pi}\hat{y} + \frac{5}{6\pi}\hat{y} = \frac{5}{3\pi}\hat{y}$$

The surface current is then found from Equation 3.6, where the unit vector pointing from region 2 into region 1 is $-\hat{z}$:

$$\bar{J}_s = \hat{n}_{21} \times \bar{H}_1 = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ 0 & 0 & -1 \\ 0 & \frac{5}{3\pi} & 0 \end{vmatrix} = \frac{5}{3\pi}\hat{x} \ \ \text{A/m}$$

Notice that $|\bar{J}_s| = 2E_i/\eta_o$ as expected.

$\lll\lll\lll\ggg\ggg\ggg$

3.3 Normal Incidence on a Dielectric

In the previous section, we considered normal incidence on a perfect conductor, which was simplified by the fact that no fields exist inside of the perfect conductor. We next consider the problem of reflection and transmission from an interface between free space and a dielectric medium. In this case, transmitted fields will exist in the second region. Assume that a plane wave propagating in free space in the z-direction is normally incident on a dielectric described by the parameters ε_2 and μ_o. Using phasor notation, we again write general expressions for the incident, reflected, and transmitted fields. Similar to the previous section, the incident electric and magnetic fields are

$$\bar{E}_i = E_i e^{-jk_o z}\hat{y} \tag{3.24}$$

$$\bar{H}_i = \frac{-E_i}{\eta_o} e^{-jk_o z}\hat{x} = H_i e^{-jk_o z}\hat{x} \tag{3.25}$$

$$\bar{E}_r = E_r e^{+jk_o z}\hat{y} \tag{3.26}$$

$$\bar{H}_r = \frac{E_r}{\eta_o} e^{+jk_o z}\hat{x} = H_r e^{+jk_o z}\hat{x} \tag{3.27}$$

The transmitted fields have the form

$$\bar{E}_t = E_t e^{-jk_2 z}\hat{y} \tag{3.28}$$

$$\bar{H}_t = -\frac{E_t}{\eta_2} e^{-jk_2 z}\hat{x} = H_t e^{-jk_2 z}\hat{x} \tag{3.29}$$

where E_t and H_t are the amplitudes of the transmitted electric and magnetic fields at $z = 0$. The transmitted fields propagate through the dielectric material with the propagation constant $k_2 = \omega\sqrt{\mu_o \varepsilon_2}$. The impedance of the dielectric material is $\eta_2 = \sqrt{\mu_o/\varepsilon_2}$. The fields are illustrated in Figure 3.4.

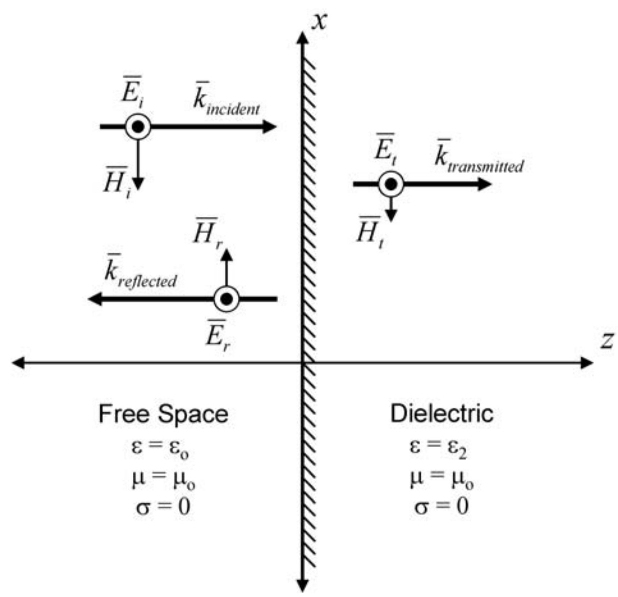


Figure 3.4: Orientation of the incident, reflected, and transmitted fields at the interface between free space and a dielectric region, at normal incidence.

We again relate the amplitudes of the reflected and transmitted fields to the incident field amplitude. We apply the boundary condition for the tangential electric fields (Equation 3.1) at the interface between the two regions, which yields

$$E_i + E_r = E_t \tag{3.30}$$

In a similar manner, we apply the boundary condition for the tangential magnetic fields (Equation 3.2) at $z = 0$, which results in

$$-\frac{E_i}{\eta_o} + \frac{E_r}{\eta_o} = -\frac{E_t}{\eta_2} \tag{3.31}$$

By combining Equations 3.30 and 3.31 we obtain an expression for the *electric field reflection coefficient* from a dielectric interface at *normal incidence* (refer to Appendix C for details):

$$\Gamma = \frac{E_r}{E_i} = \frac{\eta_2 - \eta_o}{\eta_2 + \eta_o} \tag{3.32}$$

In a similar manner, we derive the transmission coefficient to relate the transmitted electric field amplitude to the incident field amplitude. Using Equations 3.30 and 3.32 we find (refer to Appendix C for details):

$$\mathcal{T} = \frac{E_t}{E_i} = \frac{E_r + E_i}{E_i} = \Gamma + 1 = \frac{2\eta_2}{\eta_2 + \eta_o} \tag{3.33}$$

Since the two materials are lossless, the impedance is a real quantity that can be expressed in terms of the index of refraction of the material as $\eta = \eta_o/n$. In this case, the reflection and transmission coefficients can be rewritten as

$$\Gamma = \frac{\eta_o/n_2 - \eta_o}{\eta_o/n_2 + \eta_o} = \frac{1 - n_2}{1 + n_2} \tag{3.34}$$

$$\mathcal{T} = \frac{2\eta_o/n_2}{\eta_o/n_2 + \eta_o} = \frac{2}{1 + n_2} \tag{3.35}$$

Notice that for a lossless material, both the reflection and transmission coefficients are purely real quantities. Also, since the materials are lossless there is no current flow. A typical dielectric material, glass, has an index of refraction of approximately 1.5 at visible wavelengths. This means that the reflection and transmission coefficients at normal incidence from air onto a layer of glass are approximately $\Gamma = -0.2$ and $\mathcal{T} = 0.8$.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.3:

A 20 MHz plane wave in free space is normally incident on a dielectric with an index of refraction $n_2 = 1.5$. The interface between the free space and dielectric regions is the $z = 0$ plane. The electric field has an amplitude of 100 V/m and is polarized in the x-direction. Find the phasor expressions for the incident, reflected, and transmitted electric fields, and the phasor expression for the total electric field in the free space region.

Solution:

The wave vector of the incident wave is

$$\bar{k}_o = \frac{\omega}{c}\hat{z} = \frac{2\pi \times 60 \times 10^6}{3 \times 10^8}\hat{z} = 0.4\hat{z}$$

The incident electric field is

$$\bar{E}_i = E_i e^{-j\bar{k}_o \cdot \bar{r}} = 100 e^{-j0.4z}\hat{x} \ \ \text{V/m}$$

The electric field reflection coefficient from the interface is

$$\Gamma = \frac{1 - n_2}{1 + n_2} = \frac{1 - 1.5}{1 + 1.5} = -0.2$$

Since the reflected fields propagate in the -z direction, the sign of the exponent must be changed. The reflected electric field is

$$\bar{E}_r = \Gamma E_i e^{+j0.4z}\hat{x} = -20 e^{+j0.4z}\hat{x} \ \ \text{V/m}$$

The total electric field in the free space region (region 1) is found by adding the incident and reflected fields:

$$\bar{E}_1 = \bar{E}_i + \bar{E}_r = 100 e^{-j0.4z}\hat{x} - 20 e^{+j0.4z}\hat{x}$$

$$= 20\left(5e^{-j0.4z} - e^{+j0.4z}\right)\hat{x} \ \ \text{V/m}$$

The electric field transmission coefficient is

$$\mathcal{T} = \frac{2}{1 + n_2} = \frac{2}{1 + 1.5} = 0.8$$

Also, in the dielectric (region 2), the wave vector is

$$\bar{k}_2 = k_o n_2 \hat{z} = \frac{\omega}{c} n_2 \hat{z} = \frac{2\pi \times 60 \times 10^6}{3 \times 10^8} (1.5)\, \hat{z} = 0.6\hat{z}$$

The transmitted electric field is

$$\bar{E}_t = \mathcal{T} E_i e^{-j\bar{k}_2 \cdot \bar{r}} \hat{x} = 80 e^{-j0.6z} \hat{x} \text{ V/m}$$

$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

3.4 Normal Incidence on a Conducting Medium

At this point, we have considered reflection and transmission from a perfect conductor ($\sigma = \infty$) and a dielectric ($\sigma = 0$). What occurs if a wave is normally incident on a medium with a *finite, nonzero* conductivity? Assume that a plane wave propagating in free space in the z-direction is normally incident on a nonmagnetic ($\mu = \mu_o$) conductor with a permittivity ε_2 and conductivity σ_2, yielding a complex permittivity ε_c. The incident and magnetic electric and magnetic fields are again

$$\bar{E}_i = E_i e^{-jk_o z} \hat{y} \tag{3.36}$$

$$\bar{H}_i = \frac{-E_i}{\eta_o} e^{-jk_o z} \hat{x} = H_i e^{-jk_o z} \hat{x} \tag{3.37}$$

$$\bar{E}_r = E_r e^{+jk_o z} \hat{y} \tag{3.38}$$

$$\bar{H}_r = \frac{E_r}{\eta_o} e^{+jk_o z} \hat{x} = H_r e^{+jk_o z} \hat{x} \tag{3.39}$$

The transmitted fields in the conductor have the form

$$\bar{E}_t = E_t e^{-jk_c z} \hat{y} \tag{3.40}$$

$$\bar{H}_t = -\frac{E_t}{\eta_c} e^{-jk_c z} \hat{x} = H_t e^{-jk_c z} \hat{x} \tag{3.41}$$

The incident, reflected, and transmitted fields are shown in Figure 3.5. The transmitted fields propagate through the lossy conducting material with the propagation constant $k_c = \omega\sqrt{\mu_o \varepsilon_c} = \beta - j\alpha$. Because the wave vector is complex, the amplitude of the wave will decay away from the interface. The impedance of the lossy material is $\eta_c = \sqrt{\mu_o/\varepsilon_c}$, which is also a complex quantity. In Chapter 2, it was shown that as the conductivity of the medium increases, the magnitude of the impedance η_c decreases. If the medium is a perfect conductor, the impedance is zero.

Noting the similarity between this case and the case of reflection from a dielectric, we use Equation 3.32 to write an expression for the *electric field reflection coefficient* from a conductor at *normal incidence*:

$$\Gamma = \frac{E_r}{E_i} = \frac{\eta_c - \eta_o}{\eta_c + \eta_o} \tag{3.42}$$

An important point to notice is that as the conductivity increases towards infinity, the electric field reflection coefficient approaches -1:

$$\lim_{\sigma_c \to \infty} \Gamma = \lim_{\sigma_c \to \infty} (E_r/E_i) = \frac{-\eta_o}{+\eta_o} = -1 \tag{3.43}$$

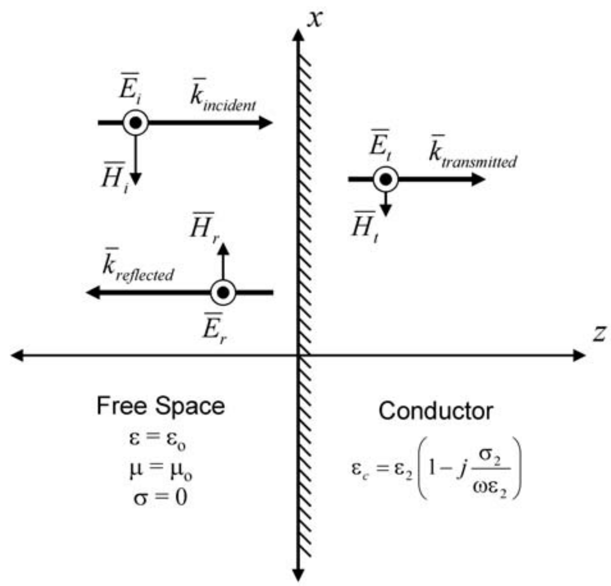


Figure 3.5: Orientation of the incident, reflected, and transmitted fields at the interface between free space and a conducting region, at normal incidence.

This agrees with the reflection coefficient we found previously for a perfect conductor.

What happens to the magnetic field when it is reflected from the conducting boundary? From Equations 3.37 and 3.39 we find that $H_i = -E_i/\eta_o$ and $H_r = E_r/\eta_o$. Using these results, we find that the magnetic field experiences a reflection

$$\frac{H_r}{H_i} = -\frac{E_r}{E_i} = -\Gamma = \frac{\eta_o - \eta_c}{\eta_c + \eta_o} \tag{3.44}$$

Notice that the magnetic field reflection approaches +1 as the conductivity increases, and is exactly +1 for a perfect conductor:

$$\lim_{\sigma_c \to \infty} (H_r/H_i) = \lim_{\sigma_c \to \infty} (-\Gamma) = \frac{+\eta_o}{+\eta_o} = +1 \tag{3.45}$$

This again agrees with the results obtained previously for a perfect conductor.

The electric field transmission coefficient from a conductor at normal incidence is

$$\mathcal{T} = \frac{E_t}{E_i} = \frac{E_r + E_i}{E_i} = \Gamma + 1 = \frac{2\eta_c}{\eta_c + \eta_o} \tag{3.46}$$

Since the magnitude of the impedance of a good conductor is much smaller than the intrinsic impedance of free space, η_o, the transmission coefficient $\mathcal{T}$ into a good conductor is small. As the conductivity approaches infinity, the transmission becomes zero, i.e. there is no electric field in a perfect conductor, as expected:

$$\lim_{\sigma_c \to \infty} \mathcal{T} = \lim_{\sigma_c \to \infty} (E_t/E_i) = \frac{0}{+\eta_o} = 0 \tag{3.47}$$

The relationship between the transmitted and incident magnetic fields is

$$\frac{H_t}{H_i} = \frac{\eta_o}{\eta_c}\frac{E_t}{E_i} = \frac{2\eta_o}{\eta_c + \eta_o} \tag{3.48}$$

Notice that as the conductivity approaches infinity, the transmission approaches $H_t/H_i = 2$. Recall that this field is dissipated by an "infinitely thin" surface current at the interface, such that the total magnetic field inside the perfect conductor is still zero.

Let us examine the current flow in region 2 in greater detail. Assume that region 2 has a finite conductivity, σ_2, and that it is a good conductor. The current density in the conducting region is then

$$\bar{J} = \sigma_2 \bar{E}_t = \sigma_2 \mathcal{T} E_i e^{-jk_c z} \hat{y} \tag{3.49}$$

The total current per unit width of the conductor measured along the x-axis is found by integrating from $0 \leq z \leq \infty$:

$$\bar{J}_{\text{tot}} = \int_0^\infty \sigma_2 \mathcal{T} E_i e^{-jk_c z} \hat{y} \;\; dz = \hat{y} \sigma_2 \mathcal{T} E_i \int_0^\infty e^{-jk_c z} dz = \frac{\sigma_2 \mathcal{T} E_i}{jk_c} \hat{y} \tag{3.50}$$

where $\bar{J}_{\text{tot}}$ has units of A/m. Assuming that the medium is a good conductor, we found in Chapter 2 that we can make the approximation $k_c = \beta - j\alpha \approx (1-j)\sqrt{\omega\mu_o\sigma_2/2}$, and therefore

$$\bar{J}_{\text{tot}} = \frac{\sigma_2 \mathcal{T} E_i}{j(1-j)\sqrt{\omega\mu_o\sigma_2/2}} \hat{y} = \frac{\mathcal{T} E_i}{(1+j)\sqrt{\omega\mu_o/(2\sigma_2)}} \hat{y} \tag{3.51}$$

Finally, if region 2 is a good conductor the complex wave impedance is $\eta_c \approx (1+j)\sqrt{\omega\mu_o/(2\sigma_2)}$, which reduces the solution to

$$\bar{J}_{\text{tot}} = \frac{1}{\eta_c} \mathcal{T} E_i \hat{y} = \frac{E_t}{\eta_c} \hat{y} = H_t \hat{y} \tag{3.52}$$

Therefore, when an electromagnetic wave is incident upon a region of *finite conductivity* the total current per unit width is equal to the tangential magnetic field at the surface. As the conductivity of the second medium approaches that of a *perfect conductor* (as $\sigma_2 \rightarrow \infty$), the magnitude of the total magnetic field at the interface becomes $|H_i + H_r| = 2E_i/\eta_o$. When the conductivity approaches infinity, the skin depth approaches zero and the total current per unit flows in an infinity thin region at the surface, i.e. becomes a surface current:

$$\lim_{\sigma_2 \rightarrow \infty} \bar{J}_{\text{tot}} = \bar{J}_s = \frac{2E_i}{\eta_o} \hat{y} \tag{3.53}$$

This current exists only at the surface of the perfect conductor. In other words, the current decays "infinitely fast." Refer to Appendix C for an alternate derivation of the surface current at the interface of a perfect conductor.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.4:

Human brain tissue has a relative permittivity and conductivity of $\varepsilon_r \approx 55$ and $\sigma \approx 2.1$ S/m, respectively, at 2.45 GHz [2]. Assume that a plane wave at 2.45 GHz carrying a time-average power of 250 mW/m^2 is normally incident from free space onto a human brain. Find the electric field reflection and transmission coefficients, Γ and $\mathcal{T}$.

Solution:

To find the reflection and transmission coefficients, we must first find the impedance of the two materials. Recall that the impedance of free space is $\eta_o \approx 377\ \Omega$. The complex permittivity of brain tissue was found in Chapter 2 to be:

$$\varepsilon_{\text{brain}} = \varepsilon_r \varepsilon_o \left(1 - j\frac{\sigma}{\omega \varepsilon_r \varepsilon_o}\right) = (4.87 - j1.36) \times 10^{-10} \quad \text{F/m}$$

From this we obtain the complex permittivity

$$\eta_{\text{brain}} = \sqrt{\frac{\mu_o}{\varepsilon_{\text{brain}}}} = 49.9 e^{j0.136} \quad \Omega$$

The reflection coefficient is

$$\Gamma = \frac{\eta_{\text{brain}} - \eta_o}{\eta_{\text{brain}} + \eta_o} = \frac{49.9e^{j0.136} - 377}{49.9e^{j0.136} + 377} = 0.768e^{j3.11}$$

From this expression, we see that the electric field will experience a 3.11 radian, or 178°, phase shift upon reflection. The transmission coefficient is

$$\mathcal{T} = \Gamma + 1 = 0.234e^{j0.120}$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.5:

Find the time-domain expression for the transmitted electric field inside the brain tissue. Assume that the electric field is polarized in the x-direction, propagates in the z-direction, and that the interface is at $z = 0$.

Solution:

We first must find the amplitude of the electric field propagating in free space. In free space ($z < 0$), the wave will have the general form:

$$\bar{E}(z < 0) = E_o e^{-jk_o z}\hat{x}$$

We can find the amplitude of the field, E_o, given the average power in the wave:

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{|\bar{E}_o|^2}{2\eta_o}\hat{z} = 0.250\hat{z} \quad \text{W/m}^2$$

From this we find that $E_o = 13.7$ V/m. The transmitted electric field in the brain tissue will have the general form

$$\bar{E}(z > 0) = \mathcal{T} E_o e^{-jk_{\text{brain}} z}\hat{x}$$

The complex propagation constant in the brain is

$$k_{\text{brain}} = \omega\sqrt{\mu_o \varepsilon_{\text{brain}}} = 384 - j52.7$$

Therefore, the phasor expression for the electric field transmitted into the brain tissue ($z > 0$) is

$$\bar{E}(z > 0) = \left(0.234e^{j0.120}\right)(13.7)\, e^{-j384z} e^{-52.7z}\hat{x}$$

$$= 3.21e^{j0.120} e^{-j384z} e^{-52.7z}\hat{x} \quad \text{V/m}$$

Notice the 0.120 radian (6.88°) phase shift occurring because of the lossy material interface. The time domain electric field is then found from the phasor expression:

$$\bar{\mathcal{E}}(z > 0, t) = 3.21\cos(1.54 \times 10^{10} t - 384z + 6.88°)e^{-52.7z}\hat{x} \quad \text{V/m}$$

The reflected field would be found in a similar manner.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

3.5 Oblique Incidence of Plane Waves at a Single Interface

3.5.1 Wave Propagation in a General Direction

Up to this point, we have studied plane waves propagating in unbounded media and waves normally incident onto an interface. In these cases, a convenient choice for the coordinate system was to assume propagation in the z-direction. Because the electromagnetic fields were normal to the direction of propagation, they were assumed to lie in the x- or y-direction. As we will see later, though, many useful electromagnetic devices rely on waves reflecting off of or transmitting through interfaces at non-normal, or "oblique" incidence. In this case, the wave vector may have components in the x-, y-, and/or z-directions. As a reminder, recall that the wave vector pointing in the direction of plane wave propagation is expressed in a Cartesian coordinate system as $\overline{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$. It is useful to keep the vector notation because in many cases the wave vector components become complex, for example if the medium is lossy. If the wave vector components are purely real, angles can be used to express the direction of propagation as shown in the Figure 3.6.

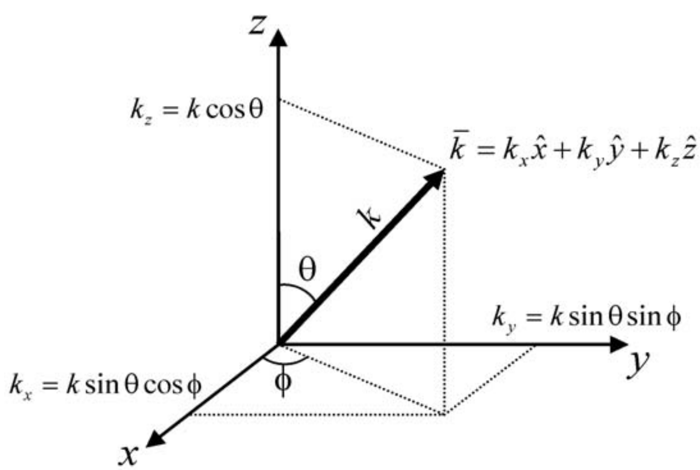


Figure 3.6: Description of wave vector in the Cartesian and spherical coordinate system. (Only when k is real can the vector be described in terms of angles.)

For simplicity, when dealing with problems of reflection and transmission off of an interface we align the coordinate axes to coincide with the interface. In addition, without loss of generality, we can often rotate the coordinate axes such that the incident wave vector is confined to the x-z or y-z plane. This greatly simplifies the analysis of the problem, as we will show in the following sections.

It is interesting to point out here that, when a plane wave is incident at an oblique angle onto an interface, not all of the power of the wave is transmitted across the interface. This can be shown by decomposing the plane wave into one component that is normal to the interface and one component that is parallel to the interface. The component that is normal to the interface will point in the direction of the boundary normal vector. If we think of the conservation of energy across the interface, only the normal components will contribute to net energy flow through the interface. We will discuss power flow across an interface in greater detail later in this chapter.

3.5.2 Plane of Incidence

Assume that a plane wave is incident onto an interface described by the $z = 0$ plane, as shown in Figure 3.7. Assume that we have aligned the coordinate axes such that the plane wave is propagating in the x-z plane at an oblique angle θ towards the planar interface. We define a unit vector, $\hat{z}$, that is normal to the boundary. The *plane of incidence* is defined as the plane that is formed by the boundary normal and the incident wave vector:

Plane of Incidence: The plane defined by the wave vector and a unit vector that is normal to the planar interface.

For the case shown in Figure 3.7, the plane of incidence is the x-z plane, since both the boundary normal and the wave vector lie in this plane.

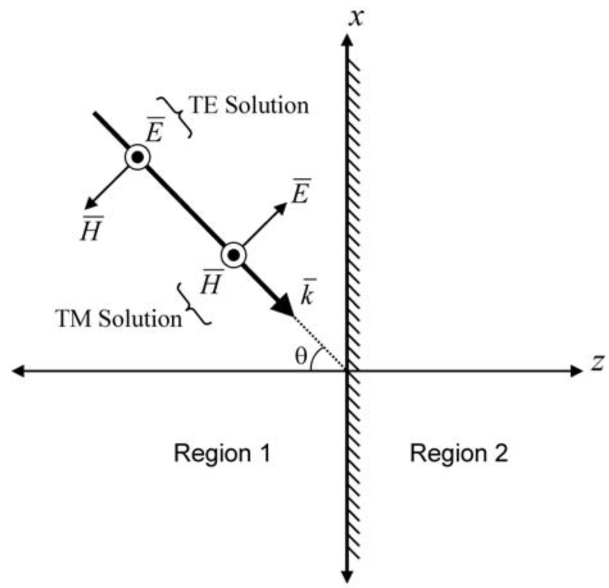


Figure 3.7: When analyzing the problem of oblique incidence onto an interface, the total solution is broken down into the transverse electric (TE) and transverse magnetic (TM) solutions. The plane of incidence is the x-z plane in this case.

3.5.3 TE and TM Solutions

Let us now look in greater detail at a plane wave that is propagating at an arbitrary angle but has a wave vector that is confined to the x-z plane, as shown in Figure 3.7. In this case, the wave vector does not have a y-component. From $\overline{k} \times \overline{E} = \omega\mu_o\overline{H}$ we obtain

$$\overline{k} \times \overline{E} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ k_x & 0 & k_z \\ E_x & E_y & E_z \end{vmatrix} = \begin{pmatrix} -k_z E_y \\ k_z E_x - k_x E_z \\ k_x E_y \end{pmatrix} = \omega\mu_o \begin{pmatrix} H_x \\ H_y \\ H_z \end{pmatrix} \tag{3.54}$$

Notice that the E_y field component can be obtained from the H_x and H_z field components. The H_y field component can be obtained from the E_x and E_z field components. In a similar manner, we can use the equation $\overline{k} \times \overline{H} = -\omega\varepsilon_o\overline{E}$ to obtain

$$\begin{pmatrix} -k_z H_y \\ k_z H_x - k_x H_z \\ k_x H_y \end{pmatrix} = -\omega\varepsilon_o \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{3.55}$$

Again we observe that the field components can be divided into two independent sets. One set is composed of the field components $\{H_x, E_y, H_z\}$, while the other set is $\{E_x, H_y, E_z\}$. These sets can be solved independently.

The first solution, $\{H_x, E_y, H_z\}$, can be completely formulated in terms of E_y and is called the *transverse electric* (TE) solution. It is called transverse electric because the electric field points in the y-direction, which is normal (or transverse) to the plane of incidence. Remember that, in this case, the plane of incidence is the x-z plane. The TE solution is also referred to as *perpendicular polarization* since the electric field is perpendicular to the plane of incidence.

The solution of the set $\{E_x, H_y, E_z\}$ can be formulated in terms of H_y and is called the *transverse magnetic* (TM) solution, because the magnetic field is transverse to the plane of incidence. The TM solution is also referred to as *parallel polarization* since the electric field is parallel to the plane of incidence.

The total solution is the sum of the TE and TM solutions. A typical problem of an arbitrary plane wave incident upon an interface would be approached as follows. First, the incident wave is decomposed into its TE and TM components. The reflected and transmitted fields are determined separately for the TE and TM cases. Then, the TE and TM reflected waves are summed to find the total reflected wave, and the TE and TM transmitted waves are summed to obtain the total transmitted wave. This separation into the independent TE and TM solutions makes the field analysis much easier.

3.5.4 Tangential Phase Matching at an Interface

In order to understand the behavior of a plane wave at an interface between two media, we must first understand the concept of "phase matching" at the interface. Consider a plane wave that is obliquely incident on a planar interface at $z = 0$, as shown in Figure 3.8. For this analysis, let us assume that two waves (a forward and backward travelling wave) will exist in both region 1 and region 2. The reason for this will become clear when we generalize the analysis to account for multiple interfaces. Here, we will focus on the transverse electric (TE) solution, which is comprised of the set of fields $\{H_x, E_y, H_z\}$ and can be formulated in terms of E_y. A similar derivation can be performed using the TM solution, but the outcome is the same. To simplify the notation, we will define $\kappa = k_x$ (the x-component of the wave vector) and $\beta = k_z$ (the z-component of the wave vector). Using this notation, a wave vector that is confined to the x-z plane will have the form

$$\bar{k} = \kappa\hat{x} + \beta\hat{z} \tag{3.56}$$

and therefore $k^2 = \kappa^2 + \beta^2$. If the wave vector is a real vector at an angle θ with respect to the interface normal, we can write $\kappa = k\sin(\theta)$ and $\beta = k\cos(\theta)$.

In the incident region, which we will call region 1, the total field is composed of two plane waves and therefore the total electric field is

$$\bar{E}_1 = \bar{E}_{1+} + \bar{E}_{1-} = \left(A_1 e^{-j\beta_{1+}z} e^{-j\kappa_{1+}x} + B_1 e^{+j\beta_{1-}z} e^{-j\kappa_{1-}x}\right)\hat{y} \tag{3.57}$$

where A_1 is the amplitude of the forward travelling electric field, and B_1 is the amplitude of the backward travelling electric field. The $+$ subscripts refer to parameters for the forward travelling wave, and the $-$ subscripts refer to parameters for the backward travelling wave. In the second region, the total electric field will be

$$\bar{E}_2 = \bar{E}_{2+} + \bar{E}_{2-} = \left(A_2 e^{-j\beta_{2+}z} e^{-j\kappa_{2+}x} + B_2 e^{+j\beta_{2-}z} e^{-j\kappa_{2-}x}\right)\hat{y} \tag{3.58}$$

Remember that, due to the boundary condition for the tangential electric fields (Equation 3.1), the tangential electric fields must be continuous across an interface. Therefore, we set the tangential electric fields in region 1 equal to the tangential electric fields in region 2 at $z = 0$ which yields:

$$A_1 e^{-j\kappa_{1+}x} + B_1 e^{-j\kappa_{1-}x} = A_2 e^{-j\kappa_{2+}x} + B_2 e^{-j\kappa_{2-}x} \tag{3.59}$$

This relationship must be valid for all values of x along the interface. The only way that this condition can hold is if the tangential components of the wave vector are identical, i.e. if $\kappa_{1+} = \kappa_{1-} = \kappa_{2+} = \kappa_{2-}$.

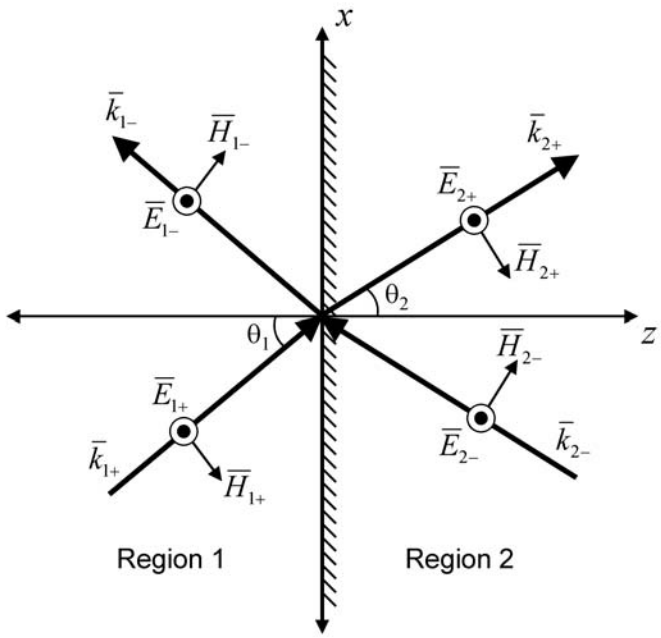


Figure 3.8: Orientation of the electromagnetic fields for the TE-polarized case. In the derivation of the field relationships it is assumed that all the electric field vectors point in the +y-direction.

Therefore, we can define the x-components of the wave vector simply as κ for the forward and backward travelling waves in *both regions*! This important conclusion is known as the *phase matching condition* at an interface:

Phase Matching Condition: At an interface, the tangential component of the wave vector is the same for all of the waves in both regions 1 and 2! This is true for both TE and TM polarizations. *This is a key conclusion with very important implications, and we will apply this condition repeatedly.*

Assuming that the tangential component of the wave vector is real, it can be written as $\kappa = k_x = \omega/v_x$, where v_x is the velocity of the wave in the x-direction. For linear media, the frequency is the same in all regions. The phase matching condition implies, then, that the velocities tangential to the interface of all the waves must be identical, which makes intuitive sense. Imagine a wave that is travelling parallel to a boundary such that part of the wave is in medium 1 and part is in medium 2. It does not make sense that the part of the wave along the interface in one material should go faster than the part in the other material!

Now that we have established that the tangential components of the wave vectors must be equal, what does this say about the normal components of the wave vectors? The magnitude of the wave vector in region 1 can be written as

$$k_1 = \sqrt{\kappa^2 + \beta_{1+}^2} = \sqrt{\kappa^2 + \beta_{1-}^2} \tag{3.60}$$

From this expression we conclude that $\beta_{1+} = \beta_{1-}$ in region 1, so we define $\beta_1 = \beta_{1+} = \beta_{1-}$. Likewise, the magnitude of the wave vector in region 2 can be written as

$$k_2 = \sqrt{\kappa^2 + \beta_{2+}^2} = \sqrt{\kappa^2 + \beta_{2-}^2} \tag{3.61}$$

It then follows that the magnitude of the normal wave vectors in the *same region* will have the *same magnitude.* Within the same region, the only difference in the wave vector for the forward and backward travelling waves is the sign of the z-component (β), which will determine whether the wave is travelling in the +z or -z direction.

Using these results, we can rewrite the expressions for the electric fields in both regions (Equations 3.57 and 3.58) for the TE case incorporating the phase matching condition:

$$\bar{E}_1 = \left(A_1 e^{-j\beta_1 z} + B_1 e^{+j\beta_1 z}\right) e^{-j\kappa x}\hat{y} \tag{3.62}$$

$$\bar{E}_2 = \left(A_2 e^{-j\beta_2 z} + B_2 e^{+j\beta_2 z}\right) e^{-j\kappa x}\hat{y} \tag{3.63}$$

3.5.5 Tangential Phase Matching: "k-Vector Surfaces" and Snell's Law

It is instructive to look at the phase matching condition for the tangential wave vectors using the "k-vector surfaces" in both regions. Considering lossless media, in region 1 the wave vector must satisfy $\bar{k}_1 \cdot \bar{k}_1 = k_1^2 = \omega^2 \mu_o \varepsilon_1$ and in region 2 it must satisfy $\bar{k}_2 \cdot \bar{k}_2 = k_2^2 = \omega^2 \mu_o \varepsilon_2$. Since the vector equation $\bar{r} \cdot \bar{r} = R^2$ describes the surface of a sphere of radius R, the wave vector equations describe spherical surfaces with radii $k_1 = \omega\sqrt{\mu_o \varepsilon_1} = k_o n_1$ in region 1 and $k_2 = \omega\sqrt{\mu_o \varepsilon_2} = k_o n_2$ in region 2. This is illustrated in Figure 3.9. Remember that the tangential wave vectors of all the waves must be the same. Also, the magnitude of the wave vectors must lie on the spherical "k-vector surface" in each region. Because the tangential component, κ, is fixed by the incident wave, and the magnitude of the wave vectors must lie on the "k-vector surface," the direction of all of the waves in each region are determined by the incident wave vector, as shown in Figure 3.9.

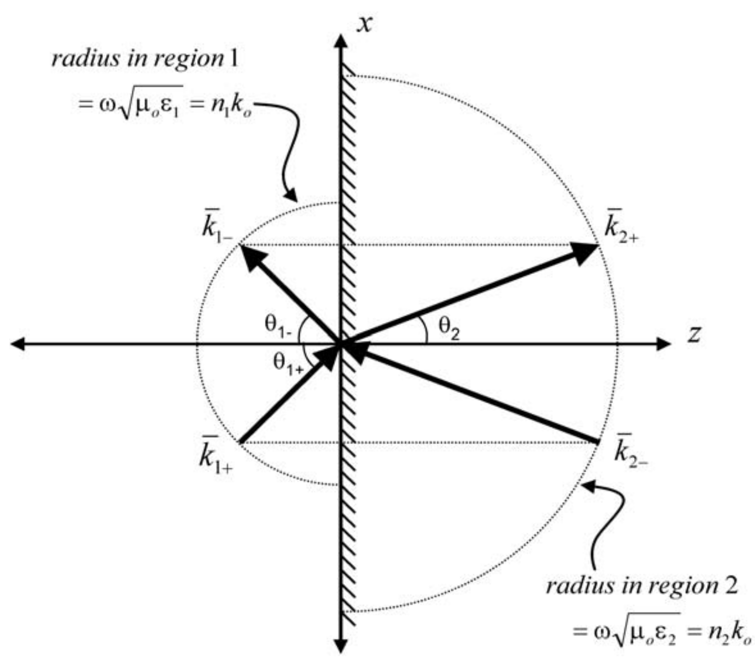


Figure 3.9: The tangential components of the wave vectors of all of the waves is the same in both regions. This determines the direction of propagation in both regions. In this example, the first (incident) region has a lower index of refraction than the second region.

Since we are assuming lossless media, the tangential component of the forward travelling wave in region 1 can be written as $\kappa = k_1 \sin(\theta_{1+})$, where θ_{1+} is the angle between the wave vector and the z-axis, as shown in Figure 3.9. Likewise, the tangential component of the backward travelling wave (the reflected wave) in region 1 can be written as $\kappa = k_1 \sin(\theta_{1-})$. Because the tangential component of the wave vector is the same, this means that $\theta_{1+} = \theta_{1-} \equiv \theta_1$. In other words, *the angle of incidence equals the angle of reflection*. This is known as *Snell's law of reflection*.

Similarly, the tangential component of the forward travelling wave in region 2 can be written as $\kappa = k_2 \sin(\theta_2)$. Equating the tangential wave vector components in regions 1 and two yields $k_1 \sin(\theta_1) = k_2 \sin(\theta_2)$. Since $k = nk_o$ in a lossless medium, this expression can be rewritten as

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \tag{3.64}$$

which is known as *Snell's law of refraction*, or often simply referred to as *Snell's law*. This is a very important conclusion, since it allows us to easily determine the angle of propagation of the transmitted wave. Note that Snell's law of refraction only applies when *both media are lossless*, and when *total internal reflection (TIR) is not occurring*. TIR will be discussed later in this chapter.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.6:

An electromagnetic wave propagating in air is incident at 30° onto a dielectric with index of refraction $n = 2.2$. Assume that the interface between the air and dielectric is the $z = 0$ plane. The wave vectors lie in the x-z plane, as shown in Figure 3.10. If the wavelength of the incident wave is $\lambda = 1.57$ μm, find the wave vectors of the incident, reflected, and transmitted waves.

Solution:

The magnitude of the wave vector in the incident (air) region is

$$k_1 = \frac{2\pi n_1}{\lambda} = \frac{2\pi}{1.57 \times 10^{-6}} = 4.00 \times 10^6 \quad \text{rad/m}$$

The components of the incident wave vector are then found from the incident angle. (Notice that the sign of each component must be determined from the direction of propagation):

$$\bar{k}_i = \kappa\hat{x} + \beta_1\hat{z} = \hat{x}k_1 \sin(\theta_1) + \hat{z}k_1 \cos(\theta_1) = 2.00 \times 10^6\hat{x} + 3.46 \times 10^6\hat{z} \quad \text{rad/m}$$

For the reflected wave, κ and β_1 remain the same. The only change is that the wave is now propagating in the *negative* z-direction:

$$\bar{k}_r = \kappa\hat{x} - \beta_1\hat{z} = 2.00 \times 10^6\hat{x} - 3.46 \times 10^6\hat{z} \quad \text{rad/m}$$

To find the wave vector of the transmitted wave, we first find the magnitude of the wave vector in the second region:

$$k_2 = \frac{2\pi n_2}{\lambda} = \frac{2\pi \times 2.2}{1.57 \times 10^{-6}} = 8.80 \times 10^6 \quad \text{rad/m}$$

In region 2, the relationship $k_2^2 = \kappa^2 + \beta_2^2$ must hold, and we know that κ is the same in both regions. Therefore, the z-component of the wave vector in region 2 is

$$\beta_2 = \sqrt{k_2^2 - \kappa^2} = \sqrt{(8.80 \times 10^6)^2 - (2.00 \times 10^6)^2} = 8.57 \times 10^6 \quad \text{rad/m}$$

Therefore, the wave vector in the second region is:

$$\bar{k}_t = \kappa\hat{x} + \beta_2\hat{z} = 2.00 \times 10^6\hat{x} + 8.57 \times 10^6\hat{z} \quad \text{rad/m}$$

Note that, since both media are lossless, we could have found the transmitted wave vector by first using Snell's law, $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$, to find the angle of the transmitted wave:

$$\theta_2 = \sin^{-1}[n_1 \sin(\theta_1)/n_2] = \sin^{-1}[\sin(30°)/2.2] = 13.1°$$

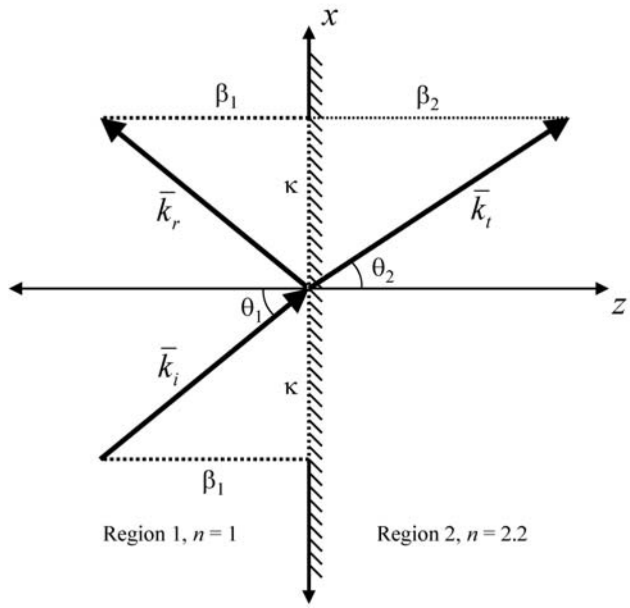


Figure 3.10: Illustration of the wave vectors for Example 3.6: a wave in air incident on a dielectric.

Using this angle, it is a simple matter to find the x- and z-components of the wave vector in region 2, since the magnitude is known.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

3.5.6 Single Interface: TE-polarized case

We now continue our solution of the TE-polarized fields to obtain an expression that relates the wave amplitudes across the dielectric interface. The general forms of the electric fields in regions 1 and 2 are given in Equations 3.62 and 3.63. We now apply the boundary condition for the tangential electric fields at the interface (Equation 3.1). By matching the tangential electric fields at the interface, i.e. setting $\bar{E}_1(z=0) = \bar{E}_2(z=0)$, it follows that the field amplitudes are related by

$$A_1 + B_1 = A_2 + B_2 \tag{3.65}$$

From the equations for the electric fields it is straightforward to solve for the magnetic fields. We find the magnetic fields from the electric fields using Faraday's law in phasor form: $\nabla \times \bar{E} = -j\omega\mu_o\bar{H}$. In region 1 we find

$$\bar{H}_1 = \frac{j}{\omega\mu_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & \bar{E}_1 & 0 \end{vmatrix} = \frac{j}{\omega\mu_o}\left(-\frac{\partial}{\partial z}\bar{E}_1\hat{x} + \frac{\partial}{\partial x}\bar{E}_1\hat{z}\right) \tag{3.66}$$

Substituting the electric field in region 1 (Equation 3.62) we find that the magnetic field in region 1 is

$$\bar{H}_1 = \frac{-\beta_1}{\omega\mu_o}\left(A_1 e^{-j\beta_1 z} - B_1 e^{+j\beta_1 z}\right) e^{-j\kappa x}\hat{x}$$

$$+\frac{\kappa}{\omega\mu_o}\left(A_1e^{-j\beta_1 z}+B_1e^{+j\beta_1 z}\right)e^{-j\kappa x}\hat{z} \tag{3.67}$$

Likewise, the magnetic field in region 2 is found from the electric field in region 2 (Equation 3.63):

$$\bar{H}_2=\frac{-\beta_2}{\omega\mu_o}\left(A_2e^{-j\beta_2 z}-B_2e^{+j\beta_2 z}\right)e^{-j\kappa x}\hat{x}$$

$$+\frac{\kappa}{\omega\mu_o}\left(A_2e^{-j\beta_2 z}+B_2e^{+j\beta_2 z}\right)e^{-j\kappa x}\hat{z} \tag{3.68}$$

The magnetic fields associated with each plane wave are shown in Figure 3.8. Recall that the tangential magnetic fields must also be continuous across the boundary (unless one of the regions is perfectly conducting, see Equation 3.2). Matching the tangential magnetic fields (the x-components) in regions 1 and 2 at $z=0$ yields

$$\beta_1A_1-\beta_1B_1=\beta_2A_2-\beta_2B_2 \tag{3.69}$$

Equations 3.65 and 3.69 can be combined into the single matrix equation

$$\begin{pmatrix}1 & 1\\ \beta_1 & -\beta_1\end{pmatrix}\begin{pmatrix}A_1\\ B_1\end{pmatrix}=\begin{pmatrix}1 & 1\\ \beta_2 & -\beta_2\end{pmatrix}\begin{pmatrix}A_2\\ B_2\end{pmatrix} \tag{3.70}$$

and solving for A_1 and B_1:

$$\begin{pmatrix}A_1\\ B_1\end{pmatrix}=\begin{pmatrix}1 & 1\\ \beta_1 & -\beta_1\end{pmatrix}^{-1}\begin{pmatrix}1 & 1\\ \beta_2 & -\beta_2\end{pmatrix}\begin{pmatrix}A_2\\ B_2\end{pmatrix} \tag{3.71}$$

Carrying out the matrix inverse and multiplication, this matrix equation simplifies to

$$\begin{pmatrix}A_1\\ B_1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1+\beta_2/\beta_1 & 1-\beta_2/\beta_1\\ 1-\beta_2/\beta_1 & 1+\beta_2/\beta_1\end{pmatrix}\begin{pmatrix}A_2\\ B_2\end{pmatrix} \tag{3.72}$$

Since we are considering only a single interface, only a transmitted (forward travelling) wave will exist in region 2. Therefore we set $B_2=0$. (As we will see later, in the case of multiple interfaces, we will always set $B=0$ in the last region.) Then for the fields in the y-direction we can calculate a reflection coefficient $r_{12}=B_1/A_1$ which relates the amplitudes of the incident and reflected electric fields, and a transmission coefficient $t_{12}=A_2/A_1$ which relates the amplitudes of the incident and transmitted electric fields. Notice that when we set $B_2=0$, Equation 3.72 will have the general form

$$\begin{pmatrix}A_1\\ B_1\end{pmatrix}=\begin{pmatrix}m_{11} & m_{12}\\ m_{21} & m_{22}\end{pmatrix}\begin{pmatrix}A_2\\ 0\end{pmatrix} \tag{3.73}$$

From this matrix equation we obtain $A_1=m_{11}A_2$ and $B_1=m_{21}A_2=(m_{21}/m_{11})A_1$. Therefore, the reflection and transmission coefficients are

$$r_{12}=\frac{B_1}{A_1}=\frac{m_{21}}{m_{11}} \tag{3.74}$$

$$t_{12}=\frac{A_2}{A_1}=\frac{1}{m_{11}} \tag{3.75}$$

By inspection, we obtain $m_{21}=(1-\beta_2/\beta_1)/2$ and $m_{11}=(1+\beta_2/\beta_1)/2$ from Equation 3.72. This results in the following expressions for the reflection and transmission coefficients in terms of the normal components of the wave vector:

$$r_{12}=\frac{E_r}{E_i}=\frac{\beta_1-\beta_2}{\beta_1+\beta_2}=\Gamma_{\mathrm{TE}} \tag{3.76}$$

$$t_{12}=\frac{E_t}{E_i}=\frac{2\beta_1}{\beta_1+\beta_2}=\mathcal{T}_{\mathrm{TE}} \tag{3.77}$$

where we have defined $E_i = A_1$ (the amplitude of the incident electric field), $E_r = B_1$ (the amplitude of the reflected electric field), and $E_t = A_2$ (the amplitude of the transmitted electric field). At this point it is important to discuss the notation we are using for the reflection and transmission coefficients. The parameters Γ and $\mathcal{T}$ are defined as the reflection and transmission coefficients *in terms of the electric field.* For the TE case, r_{12} and t_{12} relate the electric field amplitudes at the interface, and therefore $r_{12} = \Gamma_{\text{TE}}$ and $t_{12} = \mathcal{T}_{\text{TE}}$. However, when we discuss the TM case we will use the same parameters, r_{12} and t_{12}, to relate the *magnetic* fields. For the TM case, then, $r_{12} \neq \Gamma_{\text{TM}}$ and $t_{12} \neq \mathcal{T}_{\text{TM}}$. It is important to understand the difference in notation between the TE and TM cases.

Assuming lossless media, the transmission and reflection coefficients (Equations 3.76 and 3.77) can also be written in terms of angles and the material properties of the two regions:

$$r_{12} = \Gamma_{\text{TE}} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2} = \frac{\eta_2 \cos\theta_1 - \eta_1 \cos\theta_2}{\eta_2 \cos\theta_1 + \eta_1 \cos\theta_2} \tag{3.78}$$

$$t_{12} = \mathcal{T}_{\text{TE}} = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2} = \frac{2\eta_2 \cos\theta_1}{\eta_2 \cos\theta_1 + \eta_1 \cos\theta_2} \tag{3.79}$$

Notice that if we set $\theta_1 = \theta_2 = 0$, then the expressions reduce to the reflection and transmission coefficients for normal incidence. Since the total tangential electric field in region 1 must be equal to the total tangential electric field in region 2, we have $E_i + E_r = E_t$. From this it immediately follows that $1 + r_{12} = t_{12}$. This can be rewritten in terms of the electric field reflection and transmission coefficients as

$$1 + \Gamma_{\text{TE}} = \mathcal{T}_{\text{TE}} \tag{3.80}$$

Remember that this condition only holds if the TE reflection and transmission coefficients are defined in terms of the electric fields!

3.5.7 Single Interface: TM-polarized Case

We now turn our attention to the transverse magnetic (TM) solution that consists of the field components $\{E_x, H_y, E_z\}$ which may be formulated in terms of H_y. Recall that the TM polarization is also called parallel polarization since the electric field is parallel to the plane of incidence. In the TM case, we begin by formulating the solution *in terms of the magnetic fields* of the plane waves, which are oriented in the y-direction. This makes the solution for the TM case much easier since the magnetic fields only have one vector component to deal with. The magnetic fields in regions 1 and 2 are again written as the sum of forward and backward travelling waves:

$$\bar{H}_1 = \bar{H}_{1+} + \bar{H}_{1-} = \left(A_1 e^{-j\beta_1 z} + B_1 e^{+j\beta_1 z}\right) e^{-j\kappa x} \hat{y} \tag{3.81}$$

$$\bar{H}_2 = \bar{H}_{2+} + \bar{H}_{2-} = \left(A_2 e^{-j\beta_2 z} + B_2 e^{+j\beta_2 z}\right) e^{-j\kappa x} \hat{y} \tag{3.82}$$

It is very important to note that, for the TM case, A_1 and A_2 are the amplitudes of the forward travelling *magnetic* fields in regions 1 and 2, and B_1 and B_2 are the amplitudes of the backward travelling *magnetic* fields in regions 1 and 2. Also notice that in these expressions we have already applied the phase matching condition such that the tangential component of the wave vector is κ in both regions. The phase matching condition was derived for the TE case, but an analogous derivation can be performed for the TM case because the tangential magnetic field intensity is also continuous across the interface.

The electric field in both regions is found from Ampere's law in phasor form: $\nabla \times \bar{H} = j\omega\varepsilon\bar{E}$:

$$\bar{E}_1 = \frac{\beta_1}{\omega\varepsilon_1} \left(A_1 e^{-j\beta_1 z} - B_1 e^{+j\beta_1 z}\right) e^{-j\kappa x} \hat{x}$$

$$-\frac{\kappa}{\omega\varepsilon_1} \left(A_1 e^{-j\beta_1 z} + B_1 e^{+j\beta_1 z}\right) e^{-j\kappa x} \hat{z} \tag{3.83}$$

$$\bar{E}_2 = \frac{\beta_2}{\omega\varepsilon_2}\left(A_2 e^{-j\beta_2 z} - B_2 e^{+j\beta_2 z}\right) e^{-j\kappa x}\hat{x}$$

$$- \frac{\kappa}{\omega\varepsilon_2}\left(A_1 e^{-j\beta_1 z} + B_1 e^{+j\beta_1 z}\right) e^{-j\kappa x}\hat{z} \tag{3.84}$$

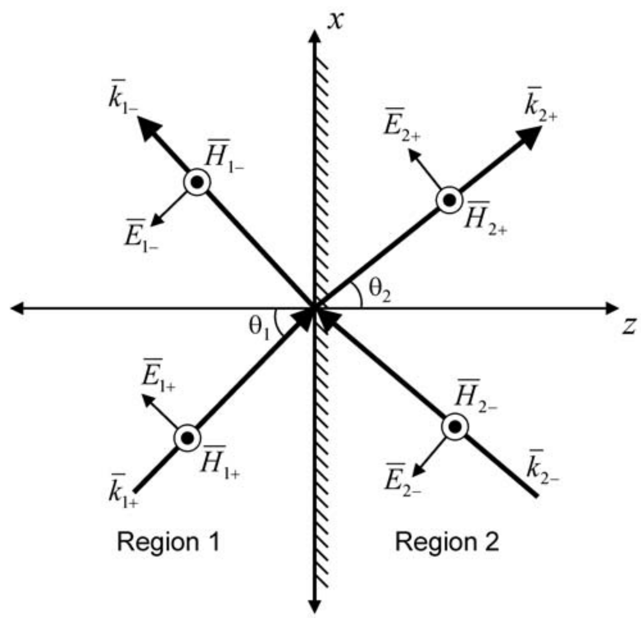


Figure 3.11: Orientation of the electromagnetic fields for the TM-polarized case. In the derivation of the field relationships it is assumed that all the magnetic field vectors point in the +y-direction.

The electric and magnetic field vectors are illustrated in Figure 3.11. Matching the tangential magnetic fields at the interface ($z = 0$) it follows that, for the TM case

$$A_1 + B_1 = A_2 + B_2 \tag{3.85}$$

Matching the tangential component (the x component) of the electric fields at the interface yields

$$\frac{\beta_1}{\varepsilon_1}A_1 - \frac{\beta_1}{\varepsilon_1}B_1 = \frac{\beta_2}{\varepsilon_2}A_2 - \frac{\beta_2}{\varepsilon_2}B_2 \tag{3.86}$$

Equations 3.85 and 3.86 can be combined into the single matrix equation:

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\delta_{12}\beta_2/\beta_1 & 1-\delta_{12}\beta_2/\beta_1 \\ 1-\delta_{12}\beta_2/\beta_1 & 1+\delta_{12}\beta_2/\beta_1 \end{pmatrix}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.87}$$

where $\delta_{12} = \varepsilon_1/\varepsilon_2$. As with the TE case, we can find a reflection coefficient $r_{12} = B_1/A_1 = m_{21}/m_{11}$ and a transmission coefficient $t_{12} = A_2/A_1 = 1/m_{11}$:

$$r_{12} = \frac{H_r}{H_i} = \frac{\varepsilon_2\beta_1 - \varepsilon_1\beta_2}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2} \tag{3.88}$$

$$t_{12} = \frac{H_t}{H_i} = \frac{2\varepsilon_2\beta_1}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2} \tag{3.89}$$

where we have defined $H_i = A_1$ (the amplitude of the incident magnetic field), $H_r = B_1$ (the amplitude of the reflected magnetic field), and $H_t = A_2$ (the amplitude of the transmitted magnetic field). If the materials are lossless, the reflection and transmission coefficients can be written in terms of angles as

$$r_{12} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2} = \frac{\eta_1 \cos\theta_1 - \eta_2 \cos\theta_2}{\eta_1 \cos\theta_1 + \eta_2 \cos\theta_2} \tag{3.90}$$

$$t_{12} = \frac{2n_2 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2} = \frac{2\eta_1 \cos\theta_1}{\eta_1 \cos\theta_1 + \eta_2 \cos\theta_2} \tag{3.91}$$

Note that from the matching of the tangential fields ($H_i + H_r = H_t$) it immediately follows that $1 + r_{12} = t_{12}$.

It is important to remember that for the TM case r_{12} and t_{12} refer to the magnetic field ratios. This means that r_{12} and t_{12} are not necessarily equal to Γ_{TM} and $\mathcal{T}_{\text{TM}}$, which are the reflection and transmission coefficients of the electric field! In some instances, however, we may desire Γ_{TM} and $\mathcal{T}_{\text{TM}}$, which are related to r_{12} and t_{12} in the following way (refer to Appendix C for details):

$$\Gamma_{\text{TM}} = \frac{E_r}{E_i} = \frac{B_1 \eta_1}{A_1 \eta_1} = r_{12} \tag{3.92}$$

$$\mathcal{T}_{\text{TM}} = \frac{E_t}{E_i} = \frac{A_2 \eta_2}{A_1 \eta_1} = \frac{\eta_2}{\eta_1} t_{12} = \frac{n_1}{n_2} t_{12} \tag{3.93}$$

The above relationships can also be derived starting with the electric field, the details of which are given in Appendix C. Remember that the *magnetic* field reflection and transmission coefficients must satisfy $1 + r_{12} = t_{12}$. This means that, for the TM case, the *electric field* reflection and transmission coefficients are related by

$$1 + \Gamma_{\text{TM}} = \frac{n_2}{n_1} \mathcal{T}_{\text{TM}} \tag{3.94}$$

Refer to Appendix C for an alternate method of deriving Equation 3.94.

For TM polarization, both the formulation in terms of the magnetic fields and the formulation in terms of the electric fields are commonly used. The formulation in terms of electric fields has the advantage that if the wave contains both TE and TM polarized fields, the formulation is consistent in terms of the electric field. In cases where we are analyzing a problem that consists entirely of TM fields, however, the analysis is significantly simplified if we use the magnetic field formulation.

3.5.8 Single Interface: Unified Formulation

By formulating the TE polarization in terms of the electric fields and the TM polarization in terms of the magnetic fields, the relationship that relates the amplitudes at an interface can be written in a unified form that holds for both TE and TM polarizations:

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \delta_{12}\beta_2/\beta_1 & 1 - \delta_{12}\beta_2/\beta_1 \\ 1 - \delta_{12}\beta_2/\beta_1 & 1 + \delta_{12}\beta_2/\beta_1 \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.95}$$

where $\delta_{12} = 1$ for TE polarization and $\delta_{12} = \varepsilon_1/\varepsilon_2 = n_1^2/n_2^2$ for TM polarization. The reflection coefficient in the unified formulation is then

$$r_{12} = \frac{\beta_1 - \delta_{12}\beta_2}{\beta_1 + \delta_{12}\beta_2} \tag{3.96}$$

and the transmission coefficient is

$$t_{12} = \frac{2\beta_1}{\beta_1 + \delta_{12}\beta_2} \tag{3.97}$$

When applying these relationships it is important to remember that, for the TE case, the reflection and transmission coefficients are with respect to the *electric* field, while for the TM case they are with respect to the *magnetic* field. If the two media are lossless, then the reflection and transmission coefficients can be rewritten in the unified formulation as

$$r_{12} = \frac{n_1 \cos\theta_1 - \delta_{12} n_2 \cos\theta_2}{n_1 \cos\theta_1 + \delta_{12} n_2 \cos\theta_2} = \frac{\eta_2 \cos\theta_1 - \delta_{12} \eta_1 \cos\theta_2}{\eta_2 \cos\theta_1 + \delta_{12} \eta_1 \cos\theta_2} \tag{3.98}$$

$$t_{12} = \frac{2 n_1 \cos\theta_1}{n_1 \cos\theta_1 + \delta_{12} n_2 \cos\theta_2} = \frac{2 \eta_2 \cos\theta_1}{\eta_2 \cos\theta_1 + \delta_{12} \eta_1 \cos\theta_2} \tag{3.99}$$

It is important to note here that, if the reflection and transmission coefficients are known, then Equation 3.95 can be written in terms of these coefficients as

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \frac{1}{t_{12}} \begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.100}$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.7:

For this example, we will show how the unified formulation can be applied to the case of oblique incidence off of a perfect conductor. Assume that a plane wave in air ($n \approx 1$) is incident upon a perfect conductor. Find expressions for the reflected electric field for the TE and TM cases.

Solution:

We begin with the TE case. From Equation 3.62, we write the incident and reflected electric fields for the TE case as

$$\bar{E}_{i,\text{TE}} = A_1 e^{-j\kappa x} e^{-j\beta_1 z} \hat{y}$$

$$\bar{E}_{r,\text{TE}} = B_1 e^{-j\kappa x} e^{+j\beta_1 z} \hat{y}$$

If region 2 is a perfect conductor, then the impedance is $\eta_2 = 0$. From Equation 3.78, the reflection coefficient for the TE case is then

$$r_{12} = \frac{B_1}{A_1} = \frac{-\eta_1 \cos\theta_2}{+\eta_1 \cos\theta_2} = -1$$

Notice that this is the same as for the case of normal incidence! Therefore we can rewrite the fields for the TE case as

$$\bar{E}_{i,\text{TE}} = E_1 e^{-j\kappa x} e^{-j\beta_1 z} \hat{y}$$

$$\bar{E}_{r,\text{TE}} = -E_1 e^{-j\kappa x} e^{+j\beta_1 z} \hat{y}$$

where E_1 is the amplitude of the incident electric field. These fields are illustrated in Figure 3.12(a). Notice that the total tangential electric field is zero at the interface. For the TM case, from Equation 3.83 we write the incident and reflected electric fields as

$$\bar{E}_{i,\text{TM}} = \frac{A_1}{\omega\varepsilon_1} (\beta_1 \hat{x} - \kappa \hat{z})\, e^{-j\kappa x} e^{-j\beta_1 z}$$

$$\bar{E}_{r,\text{TM}} = \frac{B_1}{\omega\varepsilon_1} (-\beta_1 \hat{x} - \kappa \hat{z})\, e^{-j\kappa x} e^{+j\beta_1 z}$$

From Equation 3.90 the reflection coefficient for the TM case is

$$r_{12} = \frac{B_1}{A_1} = \frac{\eta_1 \cos\theta_1}{\eta_1 \cos\theta_1} = +1$$

We can then rewrite the fields for the TM case as

$$\bar{E}_{i,\mathrm{TM}} = E_1 \left(\frac{\beta_1}{k_1}\hat{x} - \frac{\kappa}{k_1}\hat{z} \right) e^{-j\kappa x} e^{-j\beta_1 z}$$

$$\bar{E}_{r,\mathrm{TM}} = E_1 \left(-\frac{\beta_1}{k_1}\hat{x} - \frac{\kappa}{k_1}\hat{z} \right) e^{-j\kappa x} e^{+j\beta_1 z}$$

where E_1 is the amplitude of the electric field (notice that the vectors in parenthesis are unit vectors in the direction of the electric field). These fields are illustrated in Figure 3.12(b). Again we observe that the total tangential electric field is zero at the interface.

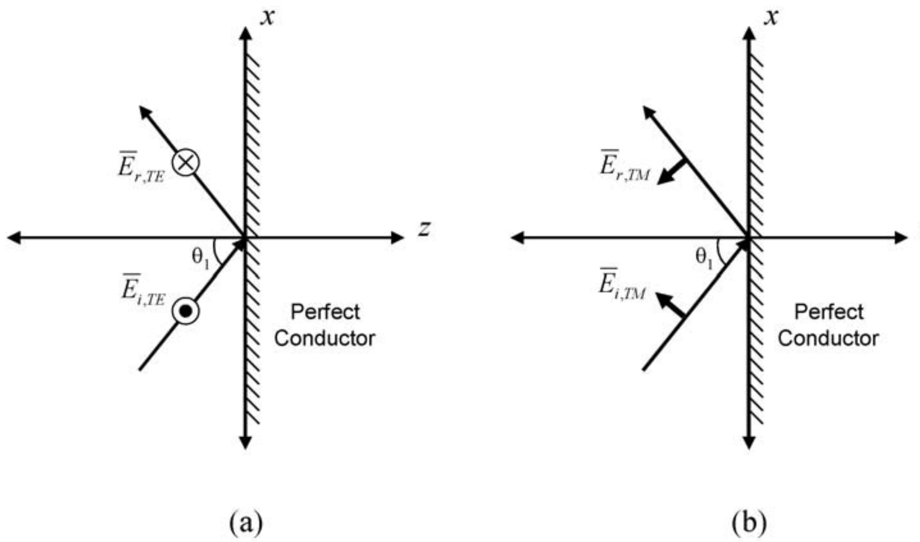


Figure 3.12: Electric field vectors for oblique incidence on a perfect conductor. (a) TE case. (b) TM case.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.8:

A plane wave in air ($n \approx 1$) is incident upon a sheet of a certain type of glass ($n \approx 1.414$). Assume that the interface between the air and glass is the $z = 0$ plane. The electric field phasor of the incident wave is

$$\bar{E}_i = E_o \left(\frac{1}{\sqrt{2}}\hat{x} + \hat{y} - \frac{1}{\sqrt{2}}\hat{z} \right) e^{-j4\times 10^6 (x+z)}$$

Find the electric field phasor of the reflected wave.

Solution:

Whenever approaching a complex electromagnetics problem such as this, it is a good idea to first make a good sketch of the fields and wave vectors! To find the reflected fields, we must find the wave vector of the reflected wave and the reflection coefficient at the interface. First, let's calculate the value of the reflected wave vector. By inspection, we see that the incident wave vector is $\bar{k}_i = 4 \times 10^6(\hat{x} + \hat{z})$. The incident angle is found from the incident wave vector components (refer to Figure 3.13):

$$\theta_1 = \tan^{-1}\left(\frac{\kappa}{\beta_1}\right) = \tan^{-1}\left(\frac{4\times 10^6}{4\times 10^6}\right) = 45^\circ$$

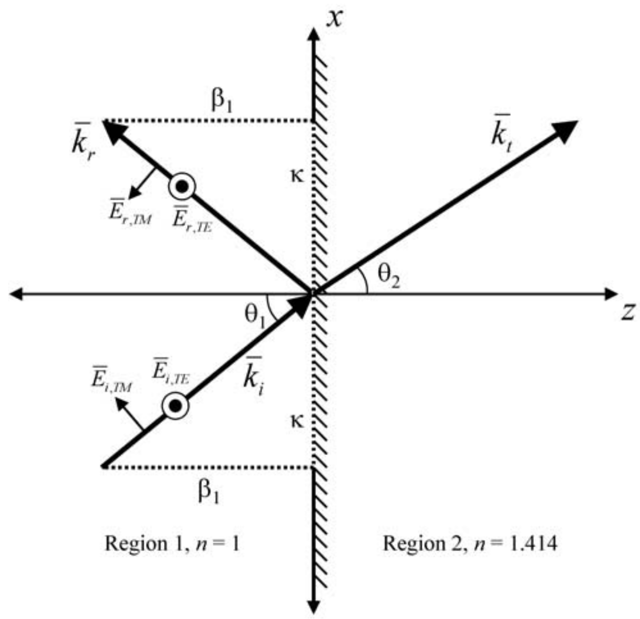


Figure 3.13: Electric fields and wave vectors for Example 3.7.

Because the reflected wave is propagating in the same medium as the incident wave and the reflected angle is the same as the incident angle, then the only change that will occur for the reflected wave vector is a change in sign of the z-component:

$$\bar{k}_r = 4 \times 10^6 \left(\hat{x} - \hat{z}\right)$$

The next step is to decide whether the incident field is TE polarized, TM polarized, or contains both TE and TM components. We identify the plane of incidence for this wave as the x-z plane. The electric field contains a component polarized in the y-direction (TE), and a component that is in the x-z plane (TM). Therefore, we must first split the field into its TE and TM components, as shown in Figure 3.13, and then find the reflected wave for each polarization separately. Then, we will add the TE and TM solutions to obtain the total field solutions for the reflected wave. By inspection, we can extract the TE and TM components:

$$\bar{E}_{i,\text{TE}} = \hat{y} E_o e^{-j4\times 10^6 (x+z)}, \quad \bar{E}_{i,\text{TM}} = \frac{E_o}{\sqrt{2}} \left(\hat{x} - \hat{z}\right) e^{-j4\times 10^6 (x+z)}$$

To find the reflection coefficient, we must know the angle of the transmitted wave, θ_2. This is found from Snell's law (Equation 3.64):

$$\theta_2 = \sin^{-1}\left(\frac{n_1}{n_2} \sin\theta_1\right) = \sin^{-1}\left(\frac{1}{1.414} \sin 45^\circ\right) = 30.0^\circ$$

The reflection coefficient for the TE polarization is then

$$\Gamma_{\text{TE}} = \frac{n_1 \cos\theta_1 - n_2 \cos\theta_2}{n_1 \cos\theta_1 + n_2 \cos\theta_2} = \frac{\cos 45^\circ - 1.414 \cos 30^\circ}{\cos 45^\circ + 1.414 \cos 30^\circ} = -0.268$$

The TE component of the reflected wave is

$$\bar{E}_{r,\text{TE}} = \Gamma_{\text{TE}} E_o e^{-j\bar{k}_r \cdot \bar{r}} \hat{y} = -0.268 E_o e^{-j4\times 10^6 (x-z)} \hat{y}$$

The TM reflection coefficient for the electric field is

$$\Gamma_{\text{TM}} = \frac{n_2 \cos\theta_1 - n_1 \cos\theta_2}{n_2 \cos\theta_1 + n_1 \cos\theta_2} = \frac{1.414 \cos 45° - \cos 30°}{1.414 \cos 45° + \cos 30°} = 0.0717$$

From Figure 3.13, we see that the x-component of the electric field vector must change sign to be consistent with Maxwell's equations. Therefore, the TM component of the reflected field is:

$$\bar{E}_{r,\text{TM}} = \Gamma_{\text{TM}} \frac{E_o}{\sqrt{2}} (-\hat{x} - \hat{z})\, e^{-j\bar{k}_r \cdot \bar{r}} = 0.0507 E_o(-\hat{x} - \hat{z}) e^{-j4\times 10^6 (x-z)}$$

Another method of obtaining the TM component is as follows. First, calculate the TM component of the incident magnetic field, $\bar{H}_{i,\text{TM}}$, which will be in the y-direction (check for yourself!). Next, calculate the *magnetic field reflection coefficient* (r_{12}) for the TM case, and use this value to calculate the reflected magnetic field, $\bar{H}_{r,\text{TM}}$, which will also be in the y-direction. Finally, calculate $\bar{E}_{r,\text{TM}}$ from $\bar{H}_{r,\text{TM}}$. This method is longer but will guarantee that the electric field vector is pointing in the correct direction.

The total reflected electric field is obtained by summing the TE and TM solutions:

$$\bar{E}_r = \bar{E}_{r,\text{TE}} + \bar{E}_{r,\text{TM}} = E_o(-0.0507\hat{x} - 0.268\hat{y} - 0.0507\hat{z}) e^{-j4\times 10^6 (x-z)}$$

≪≪≪≪≪≪≫≫≫≫≫≫

EXAMPLE 3.9:

Find the electric field phasor of the transmitted wave from Example 3.8.

Solution:

The solution process is similar to that used in Example 3.8 for the reflected field. We first must find the wave vector of the transmitted wave. Since the incident wave is in air, the magnitude must be k_o. From this we can find $k_o = 4\sqrt{2} \times 10^6$. The magnitude of the wave vector of the transmitted wave is then

$$k_t = n_2 k_o = 8 \times 10^6$$

The same result would be obtained by calculating the frequency of the wave, ω, and then $k_t = \omega\sqrt{\mu_o \varepsilon_o} n_2$. Using this result along with the transmission angle, we can find the transmitted wave vector (refer to Figure 3.14):

$$\bar{k}_t = k_t[\sin(\theta_2)\hat{x} + \cos(\theta_2)\hat{z}] = 8 \times 10^6 \left(\frac{1}{2}\hat{x} + \frac{\sqrt{3}}{2}\hat{z}\right) = 4 \times 10^6 \left(\hat{x} + \sqrt{3}\hat{z}\right)$$

where θ_2 was found in the Example 3.8 using Snell's law to be 30°.

Again, we tackle the TE and TM field solutions separately, and then add them to obtain the total electric field. The transmission coefficient for the TE component is

$$\mathcal{T}_{\text{TE}} = \frac{2n_1 \cos\theta_1}{n_1 \cos\theta_1 + n_2 \cos\theta_2} = \frac{2\cos 45°}{\cos 45° + 1.414 \cos 30°} = 0.732$$

(Notice that $1 + \Gamma_{\text{TE}} = \mathcal{T}_{\text{TE}}$ as expected.) The TE component of the transmitted electric field can easily be written now, since the polarization remains in the y-direction:

$$\bar{E}_{t,\text{TE}} = \mathcal{T}_{\text{TE}} E_o e^{-j\bar{k}_t \cdot \bar{r}} \hat{y} = 0.732 E_o e^{-j4\times 10^6 \left(x + \sqrt{3}z\right)} \hat{y}$$

The TM electric field transmission coefficient is

$$\mathcal{T}_{\text{TM}} = \frac{n_1}{n_2} t_{12} = \frac{2n_1 \cos\theta_1}{n_2 \cos\theta_1 + n_1 \cos\theta_2} = \frac{2\cos 45°}{1.414 \cos 45° + \cos 30°} = 0.758$$

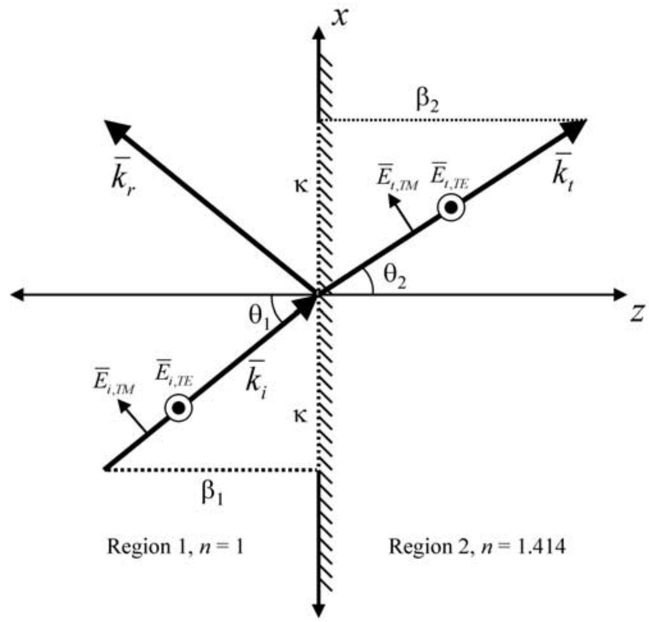


Figure 3.14: Electric fields and wave vectors for Example 3.9.

Finding the TM component of the transmitted field is a bit tricky because the magnitude and direction of the electric field vector changes in the glass. The magnitude of the transmitted field is found from the magnitude of the incident field and the transmission coefficient:

$$E_{t,\mathrm{TM}} = \mathcal{T}_{\mathrm{TM}} E_{i,\mathrm{TM}} = 0.758 \frac{E_o}{\sqrt{2}} \sqrt{1^2 + 1^2} = 0.758 E_o$$

Referring to Figure 3.14, we can then obtain the x- and z-components of the electric field, based on the angle of the transmitted wave:

$$\bar{E}_{t,\mathrm{TM}} = E_{t,\mathrm{TM}}(\hat{x}\cos\theta_2 + \hat{z}\sin\theta_2)e^{-j\bar{k}_t\cdot\bar{r}} = 0.758 E_o(\frac{\sqrt{3}}{2}\hat{x} + \frac{1}{2}\hat{z})e^{-j4\times10^6\left(x+\sqrt{3}z\right)}$$

$$= E_o(0.656\hat{x} + 0.379\hat{z})e^{-j4\times10^6\left(x+\sqrt{3}z\right)}$$

Another method of obtaining the TM component is as follows. First, calculate the TM component of the incident magnetic field, $\bar{H}_{i,\mathrm{TM}}$, which will be in the y-direction (check for yourself!). Next, calculate the *magnetic field transmission coefficient* (t_{12}) for the TM case, and use this value to calculate the transmitted magnetic field, $\bar{H}_{t,\mathrm{TM}}$, which will also be in the y-direction. Finally, calculate $\bar{E}_{t,\mathrm{TM}}$ from $\bar{H}_{t,\mathrm{TM}}$. This method is longer but will guarantee that the electric field vector is pointing in the correct direction.

The total transmitted electric field is obtained by summing the TE and TM solutions:

$$\bar{E}_t = \bar{E}_{t,\mathrm{TE}} + \bar{E}_{t,\mathrm{TM}} = E_o(0.656\hat{x} + 0.732\hat{y} + 0.379\hat{z})e^{-j4\times10^6\left(x+\sqrt{3}z\right)}$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

3.5.9 Total Internal Reflection

Total internal reflection (TIR) occurs at the interface between two dielectric (lossless) media when the incident wave is completely reflected at the interface. TIR can only occur if the *second region has a lower index of refraction than the incident region*, and if the angle of incidence is greater than a certain angle, called the critical angle. One easy way to remember this is to think of being in a swimming pool under the water with your eyes open. If you look straight up through the water/air interface, you will be able to see the sky above. As you look at the surface at an increasing angle, though, eventually you will not be able to see out of the water anymore. Instead, the surface will look like a mirror and you will only see a reflection of the bottom of the pool. This does not happen if you are in the air and looking into the water. This phenomenon occurs because the water has a higher index of refraction than the outside air.

The "k-vector surface" at the interface between a high index of refraction material and a lower index of refraction material is shown in Figure 3.15. As always, the tangential components of the incident, reflected, and transmitted wave vectors must be equal. As the angle of incidence increases, there will come a point at which the tangential component of the wave vector will be equal to k_2, so the normal component of the wave vector in the second region must be zero. This special angle is called the critical angle, θ_c, and occurs when $k_1 \sin\theta_c = k_2$, which can be rewritten as

$$\sin\theta_c = \frac{n_2}{n_1} \tag{3.101}$$

The wave vector in the second region, $\bar{k}_2 = \kappa\hat{x} + \beta_2\hat{z}$, must satisfy the relationship $k_2^2 = \kappa^2 + \beta_2^2$. At angles greater than θ_c, the tangential component, κ, is *larger* than the magnitude of the wave vector in the second region, k_2. Therefore, for the relationship to hold, β_2^2 *must be negative*, which implies the normal component of the wave vector in the second region is an imaginary quantity. Remember that an imaginary propagation constant signifies an attenuating, not propagating, wave. This can be observed if we define $\sqrt{\beta_2^2} = -j\alpha_2$. Then, the wave attenuates in the z-direction since $e^{-j(-j\alpha_2)z} = e^{-\alpha_2 z}$. Because there is no propagating wave in the second region, TIR occurs. Note, however, that this does *not* mean that there is no power flowing in the second region. It only implies that no power is leaving the interface, i.e. all of the power is flowing parallel to the interface.

The reflection coefficient for the case of TIR can be found from the general form of the reflection coefficient (Equation 3.96). Remember that TIR can only occur when $n_1 > n_2$ and $\kappa > k_o n_2$. Since $\kappa^2 + \beta_2^2 = (k_o n_2)^2$ in the second region and β_2^2 is negative, then for the TIR case we can write the normal component of the wave vector in the second region as

$$\beta_2 = \pm j\sqrt{\kappa^2 - k_o^2 n_2^2} \equiv \pm j\alpha_2 \tag{3.102}$$

The sign must be chosen such that wave decays away from the interface. For this reason we need to select the negative sign so that $e^{-j(-j\alpha_2)z} = e^{-\alpha_2 z}$. If we incorrectly select the positive sign, we would have $e^{-j(+j\alpha_2)z} = e^{+\alpha_2 z}$, which represents a wave that is exponentially *growing* away from the interface. From Equation 3.96, the reflection coefficient for the case of TIR is

$$r_{12} = \frac{\beta_1 + j\delta_{12}\alpha_2}{\beta_1 - j\delta_{12}\alpha_2} = e^{j2\phi} \tag{3.103}$$

where $\phi = \tan^{-1}(\delta_{12}\alpha_2/\beta_1)$, and $\delta_{12} = 1$ for the TE case and $\delta_{12} = \varepsilon_1/\varepsilon_2 = n_1^2/n_2^2$ for the TM case. Also, remember that r_{12} is defined in terms of the electric fields for the TE case, and in terms of the magnetic fields for the TM case. Notice that while the critical angle and the magnitude of the reflection coefficient are independent of polarization, the phase shift, ϕ, experienced by a wave undergoing TIR *is polarization dependent*.

3.5.10 Brewster's Angle

Brewster's angle is a special angle at which waves are *not reflected* from an interface. Mathematically speaking, it is the angle at which $r_{12} = 0$. Assume that a plane wave is incident on a single interface at an

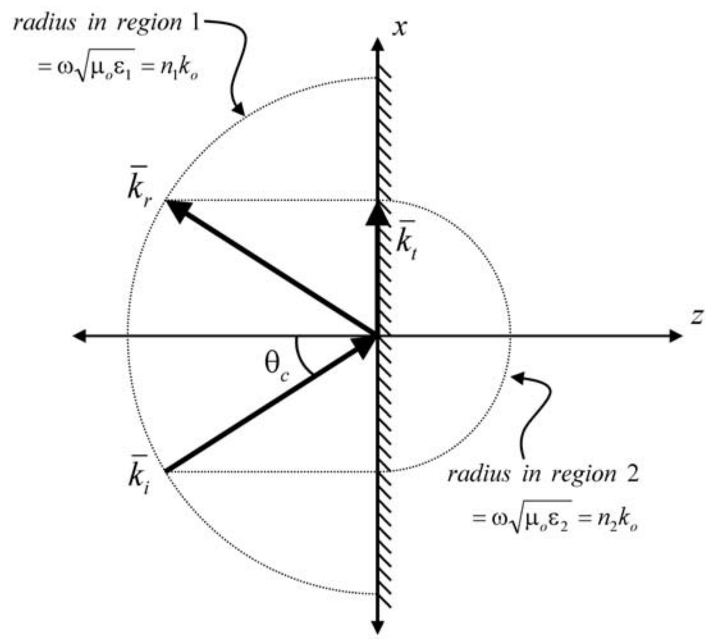


Figure 3.15: The onset of total internal refraction (TIR). TIR occurs when the index of refraction of the first region is larger than the second (think of sitting under water in a swimming pool) and the incident angle is greater than the critical angle, θ_c. Since the tangential components of the incident, reflected and transmitted waves must be equal, the normal component of the wave vector in the second region is zero at the critical angle. At angles greater than the critical angle, $\kappa > k_2$ and the normal component of the wave vector in the second region is an imaginary quantity.

angle that is less than the critical angle, θ_c. In this case, there will be propagating waves in both regions 1 and 2. From Equation 3.96 we see that the reflection coefficient will be equal to zero if $\beta_1 = \delta_{12}\beta_2$. For the TE case, this would require that $\beta_1 = \beta_2$ which can only occur in the trivial case in which $k_1 = k_2$. This means that the material properties are identical on both sides of the interface and there is no interface! However, for the TM case the reflection coefficient is zero if $\varepsilon_2\beta_1 = \varepsilon_1\beta_2$. If both regions 1 and 2 are lossless, then this condition occurs at Brewster's angle, θ_B, which is given by

$$\theta_B = \tan^{-1}\sqrt{\frac{\varepsilon_2}{\varepsilon_1}} = \tan^{-1}\frac{n_2}{n_1} \tag{3.104}$$

Remember that Brewster's angle occurs *only for TM waves in lossless media.* Refer to Appendix C for a derivation of Brewster's angle.

3.5.11 Example: Reflection from an Air/Glass Interface

Consider an electromagnetic wave propagating through air ($n \approx 1$). The wave is incident on a boundary between air and glass, which is assumed to have an index of refraction $n = 1.5$. Let us examine what happens to the magnitude of the electric field reflection coefficient, Γ, for the TE and TM cases as the incident angle, θ_1 is varied. Figure 3.16(a) illustrates the magnitude of the electric field reflection coefficients for this case. At normal incidence ($\theta_1 = 0°$), the reflection coefficients for the TE and TM cases are equal, as expected. As the incident angle is increased, Brewster's angle is clearly evident for the TM wave. Also, notice that since the index of refraction of region 1 is lower than region 2, no TIR is observed.

Now consider the opposite case, in which a wave is propagating in glass and is incident on a glass/air boundary. This case is illustrated in Figure 3.16(b). Since now region 1 has a greater index of refraction than region 2, we observe a critical angle at which the magnitude of the reflection coefficients for both TE and TM polarizations become unity. Note that these graphs show the magnitude only; there is also a varying phase shift that occurs upon reflection depending upon the incident angle.

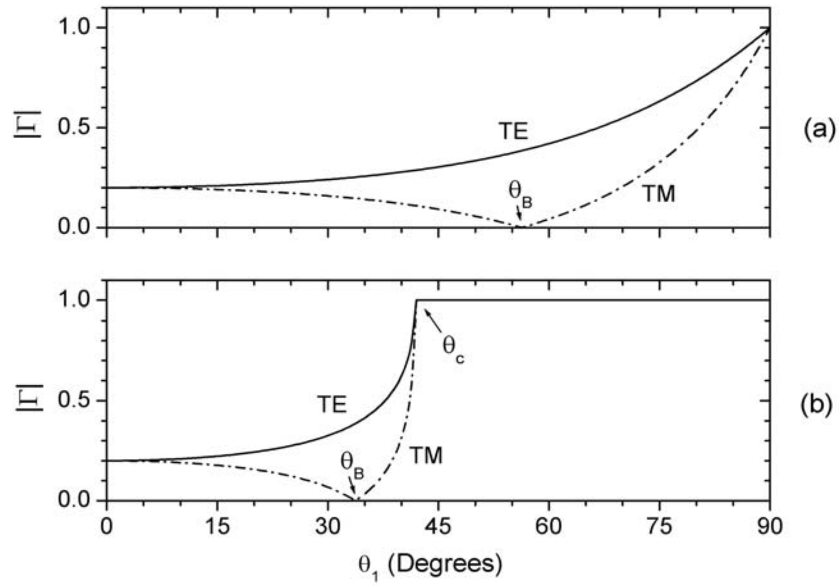


Figure 3.16: Magnitude of the electric field reflection coefficient, Γ, for varying angles of incidence, θ_1. (a) In this case, $n_1 = 1$ and $n_2 = 1.5$. Brewster's angle (θ_B) is clearly visible for the TM wave at approximately 56°. (b) In this case, $n_1 = 1.5$ and $n_2 = 1$. Since $n_1 > n_2$, there is total internal reflection past the critical angle, θ_c, which occurs at approximately 42°. Brewster's angle (θ_B) is observed for the TM wave at approximately 34°.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.10:

Assume that an electromagnetic wave propagating in glass ($n_1 = 1.414$) is incident upon a glass/air boundary, as shown in Figure 3.17. The boundary is the $z = 0$ plane. Find the wave vector in regions 1 and 2 if the electromagnetic wave is incident onto the interface at $\theta_1 = 75°$. Assume that the free space wavelength of the incident wave is $\lambda_o = 1.57$ μm.

Solution:

In this problem, $n_1 > n_2$, which means that TIR may be occurring. The critical angle for this interface is

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) = \sin^{-1}\left(\frac{1}{1.414}\right) = 45°$$

Since the incident angle is greater than the critical angle, we will have TIR. If TIR were *not* occurring, we could use Snell's law to calculate the angle of the transmitted wave vector, and then find the vector components. However, when TIR occurs, Snell's law no longer gives a valid angle in the second region because one of the wave vector components in the second region is imaginary! (Check this for yourself.) For this reason, we deal with the wave vector components directly, and apply the fact that the tangential wave vector components must be equal in both regions. The magnitude of the wave vector in region 1 is

$$k_i = k_1 = n_1 \left(\frac{2\pi}{\lambda_o}\right) = 1.414 \left(\frac{2\pi}{1.57 \times 10^{-6}}\right) = 5.66 \times 10^6 \quad \text{rad/m}$$

We calculate the wave vector in region 1 from the incident angle:

$$\bar{k}_1 = \kappa\hat{x} + \beta_1\hat{z} = k_1 \sin\theta_1\hat{x} + k_1 \cos\theta_1\hat{z} = 5.47 \times 10^6\hat{x} + 1.46 \times 10^6\hat{z}$$

In region 2, the tangential component (κ) must be the same. Also, the magnitude of the wave vector is

$$k_2 = n_2 \left(\frac{2\pi}{\lambda_o}\right) = \left(\frac{2\pi}{1.57 \times 10^{-6}}\right) = 4.00 \times 10^6 \quad \text{rad/m}$$

Therefore, the z-component of the wave vector must be

$$\beta_2 = \sqrt{k_2^2 - \kappa^2} = \sqrt{(4.00 \times 10^6)^2 - (5.47 \times 10^6)^2} = -j3.73 \times 10^6 = -j\alpha_2$$

Notice that we have chosen the negative root in order to obtain a wave that decays *away* from the interface. The wave vector in region 2 is then

$$\bar{k}_2 = \kappa\hat{x} + \beta_2\hat{z} = 5.47 \times 10^6\hat{x} - j3.73 \times 10^6\hat{z}$$

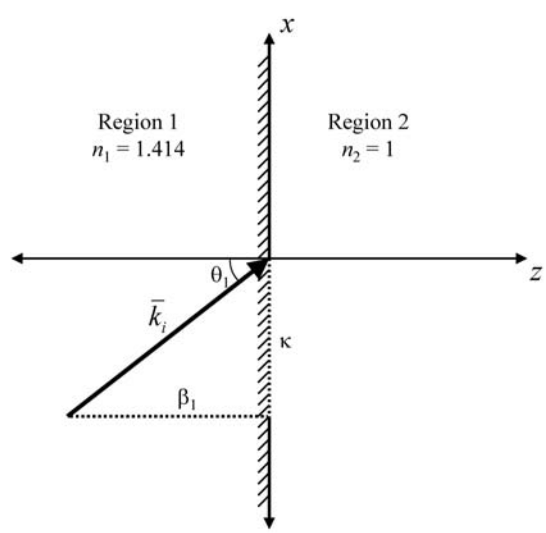


Figure 3.17: Glass/air interface described in Example 3.9.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 3.11:

For the air/glass interface described in Example 3.8 ($n_1 = 1$, $n_2 = 1.414$), find the incident angle required such that only the TE component of the wave is reflected from the interface. What is the TM electric field transmission coefficient ($\mathcal{T}_{\mathrm{TM}}$) at this angle?

Solution:

If the wave is incident at Brewster's angle, the TM reflection coefficient will be zero, and only the TE component will be reflected. From Equation 3.104, we find that Brewster's angle is

$$\theta_B = \tan^{-1}\frac{n_2}{n_1} = \tan^{-1}\frac{1.414}{1} = 54.7^\circ = \theta_1$$

Using Snell's law, the transmitted angle is found to be $\theta_2 = 35.3^\circ$. Just as a check, let's calculate the electric field reflection coefficient at $\theta_1 = \theta_B$:

$$\Gamma_{\text{TM}} = \frac{n_2\cos\theta_1 - n_1\cos\theta_2}{n_2\cos\theta_1 + n_1\cos\theta_2} = \frac{1.414\cos 54.7^\circ - \cos 35.3^\circ}{1.414\cos 54.7^\circ + \cos 35.3^\circ} = 0$$

The electric field transmission coefficient for the TM case is

$$\mathcal{T}_{\text{TM}} = \frac{n_1}{n_2}t_{12} = \frac{2n_1\cos\theta_1}{n_2\cos\theta_1 + n_1\cos\theta_2} = \frac{2\cos 54.7^\circ}{1.414\cos 54.7^\circ + \cos 35.3^\circ} = 0.71$$

This result can be confusing at first! If 100% transmission is occurring, then why is $\mathcal{T}_{\text{TM}} = 0.71$? The reason is because the electromagnetic *power* must be conserved across the interface, and not necessarily the magnitude of the *electric field*. It is this topic that will be addressed in the next section.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

3.6 Power Flow at an Interface

We next turn our attention to the flow of power that occurs at an interface when a plane wave is incident at an oblique angle. It makes intuitive sense that the power should be conserved across the interface, since power is not created or absorbed by the interface. But, when the wave is at an oblique angle of incidence, what power are we referring to? If a plane wave is incident at an angle, the wave can be decomposed into one part that is propagating towards the interface and another part that is propagating parallel to the interface. Power conservation only refers to the components that propagate normal to (i.e. towards or away from) the interface.

3.6.1 TE Polarization

To illustrate the conservation of power across an interface, we will calculate time-average power flow per unit area, $\langle\bar{\mathcal{S}}\rangle$, of the incident, reflected, and transmitted waves at the interface ($z = 0$). In this case, we will assume that both media are lossless. For the TE case, the electric field phasor of the incident wave is $\bar{E}_i = \hat{y}E_i e^{-j(\kappa x + \beta_1 z)}$. Therefore, the incident magnetic field phasor is

$$\bar{H}_i = \frac{\bar{k}_i \times \bar{E}_i}{\omega\mu_o} = \frac{1}{\omega\mu_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \kappa & 0 & \beta_1 \\ 0 & E_i e^{-j(\kappa x+\beta_1 z)} & 0 \end{vmatrix} = \frac{E_i}{\omega\mu_o}\left(-\beta_1\hat{x} + \kappa\hat{z}\right)e^{-j(\kappa x+\beta_1 z)} \tag{3.105}$$

The time-average power of the incident wave is then

$$\left\langle \bar{\mathcal{S}}_i \right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{E}_i \times \bar{H}_i^{\,*}\right\}$$

$$= \frac{1}{2\omega\mu_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ 0 & E_i e^{-j(\kappa x+\beta_1 z)} & 0 \\ -\beta_1 E_i^* e^{+j(\kappa x+\beta_1 z)} & 0 & \kappa E_i^* e^{+j(\kappa x+\beta_1 z)} \end{vmatrix}$$

$$= \frac{|E_i|^2}{2\omega\mu_o} \mathcal{R}e \{\kappa\hat{x} + \beta_1\hat{z}\} \tag{3.106}$$

Note that we have used $E_i E_i^* = |E_i|^2$ in deriving this relationship, and have assumed that $\kappa^* = \kappa$ and $\beta_1^* = \beta_1$ since the medium is lossless and therefore κ and β_1 are real.

For the reflected wave, the electric field amplitude is $E_r = \Gamma_{\mathrm{TE}} E_i$, and the wave propagates away from the interface with a wave vector $\bar{k}_r = \kappa\hat{x} - \beta_1\hat{z}$. Using the same procedure as the incident wave, the time-average power of the reflected wave is

$$\left\langle \bar{\mathcal{S}}_r \right\rangle = \frac{|E_i|^2 |\Gamma_{\mathrm{TE}}|^2}{2\omega\mu_o} \mathcal{R}e \{\kappa\hat{x} - \beta_1\hat{z}\} \tag{3.107}$$

The amplitude of the transmitted electric field is $E_t = \mathcal{T}_{\mathrm{TE}} E_i$. In this case, we can not assume that $\beta_2^* = \beta_2$ since the wave in the second region may be decaying (because of TIR) which would lead to an imaginary value of β_2. Therefore, the time-averaged power of the transmitted wave at the interface is

$$\left\langle \bar{\mathcal{S}}_t \right\rangle = \frac{|E_i|^2 |\mathcal{T}_{\mathrm{TE}}|^2}{2\omega\mu_o} \mathcal{R}e \{\kappa\hat{x} + \beta_2^*\hat{z}\} \tag{3.108}$$

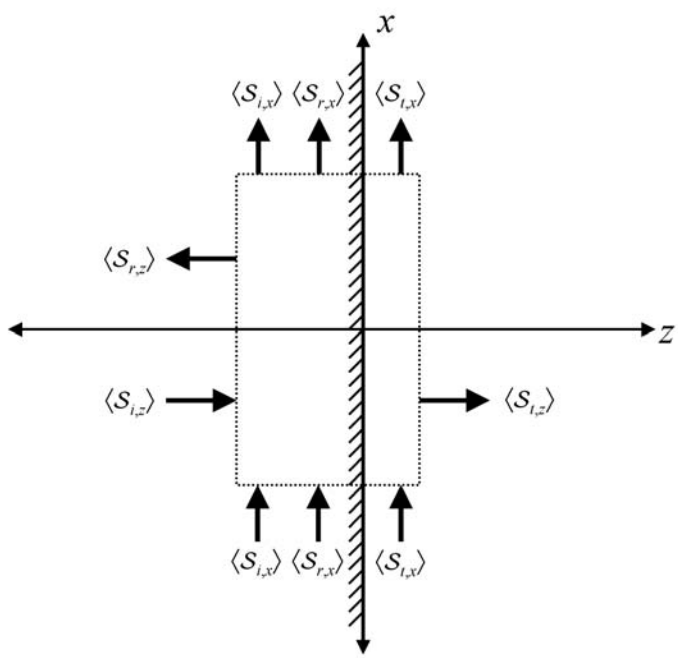


Figure 3.18: Power is conserved at the interface if the total power flowing into a volume is equal to the total power flowing out of the same volume.

To check power conservation at the interface, we choose a sample volume across the boundary as shown in Figure 3.18. For power to be conserved, the total time-average power flowing into the volume must be equal to the total time-average power flowing out of the volume. Obviously, the x-directed components are conserved since they simply flow along the interface. Therefore, we are left with the z-components. Let us denote the cross-sectional area of the sample volume in the x-y plane as A. Since $\langle \bar{\mathcal{S}} \rangle$ is the time-average power flow *per unit area*, the total time-average power of the incident, reflected, and transmitted waves in the z-direction is:

$$P_{i,z} = A \langle \mathcal{S}_{i,z} \rangle \tag{3.109}$$

$$P_{r,z} = A \langle \mathcal{S}_{r,z} \rangle \tag{3.110}$$

$$P_{t,z} = A \langle \mathcal{S}_{t,z} \rangle \tag{3.111}$$

Setting $P_{i,z} + P_{r,z} = P_{t,z}$ yields

$$\beta_1 - \beta_1 |\Gamma_{\text{TE}}|^2 = \mathcal{R}e\left\{\beta_2^*\right\} |\mathcal{T}_{\text{TE}}|^2 \tag{3.112}$$

By substituting the expressions for Γ_{TE} and $\mathcal{T}_{\text{TE}}$ (Equations 3.76 and 3.77), it is easily shown that power is indeed conserved across the interface (refer to Appendix C for details).

3.6.2 TM Polarization

Following an analogous derivation it follows that for TM-polarized waves at oblique incidence, the time-average power of the incident, reflected, and transmitted waves at the interface is

$$\left\langle \bar{\mathcal{S}}_i \right\rangle = \frac{|H_i|^2}{2\omega\varepsilon_1} \mathcal{R}e\left\{\kappa\hat{x} + \beta_1\hat{z}\right\} \tag{3.113}$$

$$\left\langle \bar{\mathcal{S}}_r \right\rangle = \frac{|H_i|^2 |\Gamma_{\text{TM}}|^2}{2\omega\varepsilon_1} \mathcal{R}e\left\{\kappa\hat{x} - \beta_1\hat{z}\right\} \tag{3.114}$$

$$\left\langle \bar{\mathcal{S}}_t \right\rangle = \frac{|H_i|^2 |\mathcal{T}_{\text{TM}}|^2}{2\omega\varepsilon_1} \mathcal{R}e\left\{\kappa\hat{x} + \beta_2\hat{z}\right\} \tag{3.115}$$

Note that the permittivity in the denominator is ε_1 in all three cases. Setting the total, z-directed, time-average power in region 1 equal to the z-directed time-average power in region 2 at the interface yields

$$\beta_1 - \beta_1 |\Gamma_{\text{TM}}|^2 = \mathcal{R}e\left\{\beta_2\right\} |\mathcal{T}_{\text{TM}}|^2 \tag{3.116}$$

By substituting the expressions for Γ_{TM} and $\mathcal{T}_{\text{TM}}$ (Equations 3.92 and 3.93), it is easily shown that power is indeed conserved across the interface (refer to Appendix C for details).

3.6.3 Reflectance and Transmittance

We have previously defined reflection and transmission coefficients at the interface, Γ and $\mathcal{T}$, in terms of the ratios of the electric fields. In addition, we can define coefficients in terms of the ratios of the incident, reflected, and transmitted *power*. These coefficients are called the *reflectance* (R) and the *transmittance* (T), and are defined as

$$R = \frac{P_{r,z}}{P_{i,z}} = \frac{\langle \mathcal{S}_{r,z} \rangle}{\langle \mathcal{S}_{i,z} \rangle} \tag{3.117}$$

$$T = \frac{P_{t,z}}{P_{i,z}} = \frac{\langle \mathcal{S}_{t,z} \rangle}{\langle \mathcal{S}_{i,z} \rangle} \tag{3.118}$$

For both the TE and TM polarizations, it can be shown that

$$R = |\Gamma|^2 \tag{3.119}$$

where $\Gamma = \Gamma_{\text{TE}}$ for TE polarization or $\Gamma = \Gamma_{\text{TM}}$ for TM polarization. The transmittance is

$$T = |\mathcal{T}|^2 \mathcal{R}e\left\{\frac{\beta_2^*}{\beta_1}\right\} \tag{3.120}$$

where $\mathcal{T} = \mathcal{T}_{\text{TE}}$ for TE polarization or $\mathcal{T} = \mathcal{T}_{\text{TM}}$ for TM polarization. In terms of the unified formulation presented earlier in this chapter, the reflectance and transmittance can be written as

$$R = |r_{12}|^2 \tag{3.121}$$

$$T = |t_{12}|^2 \mathcal{R}e \left\{ \frac{\beta_2^*}{\beta_1} \delta_{12} \right\} \tag{3.122}$$

where $\delta_{12} = 1$ for TE polarization, and $\delta_{12} = \varepsilon_1/\varepsilon_2$ for TM-polarization. Remember that in the unified formulation, the reflection and transmission coefficients r_{12} and t_{12} are in terms of the electric fields for TE polarization and in terms of the magnetic fields for TM polarization!

3.7 Standing Waves

An interesting phenomenon that occurs when a wave is reflected from an interface is the formation of a "standing wave," which is a portion of the wave that is fixed in space and does not propagate. Consider a TE polarized plane wave that is propagating in a medium with index of refraction n_1 and is reflected from an interface (an analogous argument can be presented for the TM case). The incident wave is at an angle θ_1 with respect to the interface normal. The reflected wave is then also at an angle θ_1 with respect to the normal. The two plane waves in region 1 can be written as

$$\bar{E}_i = E_i e^{-j\beta_1 z} e^{-j\kappa x} \hat{y} \tag{3.123}$$

$$\bar{E}_r = E_r e^{+j\beta_1 z} e^{-j\kappa x} \hat{y} = \Gamma_{\text{TE}} E_i e^{+j\beta_1 z} e^{-j\kappa x} \hat{y} \tag{3.124}$$

Notice that the maximum amplitude of the incident wave is $|\bar{E}_i| = E_i$ and the maximum amplitude of the reflected wave is $|\bar{E}_r| = \Gamma_{\text{TE}} E_i$, and that these amplitudes are independent of the position along the z-axis. The total electric field in region 1 is then

$$\begin{aligned} \bar{E}_{\text{tot}} = \bar{E}_i + \bar{E}_r &= E_i e^{-j\beta_1 z} e^{-j\kappa x} \hat{y} + E_r e^{+j\beta_1 z} e^{-j\kappa x} \hat{y} \\ &= E_i \left(e^{-j\beta_1 z} + |\Gamma_{\text{TE}}| e^{j\phi} e^{+j\beta_1 z} \right) e^{-j\kappa x} \hat{y} \\ &= E_i e^{-j(\kappa x + \beta_1 z)} [1 + |\Gamma_{\text{TE}}| e^{j(2\beta_1 z + \phi)}] \hat{y} \end{aligned} \tag{3.125}$$

Note that we have written the complex reflection coefficient in polar form as $\Gamma_{\text{TE}} = |\Gamma_{\text{TE}}| e^{j\phi}$. Calculating the maximum amplitude of the combined electric field in region 1, we find

$$|\bar{E}_{\text{tot}}| = |E_i e^{-j(\kappa x + \beta_1 z)}| \cdot |1 + |\Gamma_{\text{TE}}| e^{j(2\beta_1 z + \phi)}| \tag{3.126}$$

The magnitude of the first term is E_i, while the magnitude of the second term is

$$|1 + |\Gamma_{\text{TE}}| e^{j(2\beta_1 z + \phi)}| = |1 + |\Gamma_{\text{TE}}| \cos(2\beta_1 z + \phi) + j |\Gamma_{\text{TE}}| \sin(2\beta_1 z + \phi)| \tag{3.127}$$

Notice that the amplitude of the ***total field*** is no longer a constant value, but instead is a periodic function in the z-direction! The maximum value that the wave can assume at any position is called the *standing wave envelope*. The behavior of the standing wave envelope is most easily visualized using a graphical method called the "crank diagram." The crank diagram is simply a plot of Equation 3.127 in the complex plane for increasing values of z. At $z = 0$, the standing wave envelope is proportional to the length of a line extending from the origin to the point "A" shown in Figure 3.19(a). As the distance from the interface is increased (z becomes increasingly more negative), the phase change causes the "crank" to rotate in a clockwise direction, as shown in Figure 3.19(b).

From the crank diagram, it is easily seen that the standing wave envelope has a maximum amplitude of $E_i(1 + |\Gamma_{\text{TE}}|)$ and a minimum amplitude of $E_i(1 - |\Gamma_{\text{TE}}|)$. Although not presented here, we would obtain a similar solution for TM polarized fields. The *standing wave ratio* (SWR) is the ratio of the maximum to minimum amplitude of the standing wave envelope:

$$\text{SWR} = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{3.128}$$

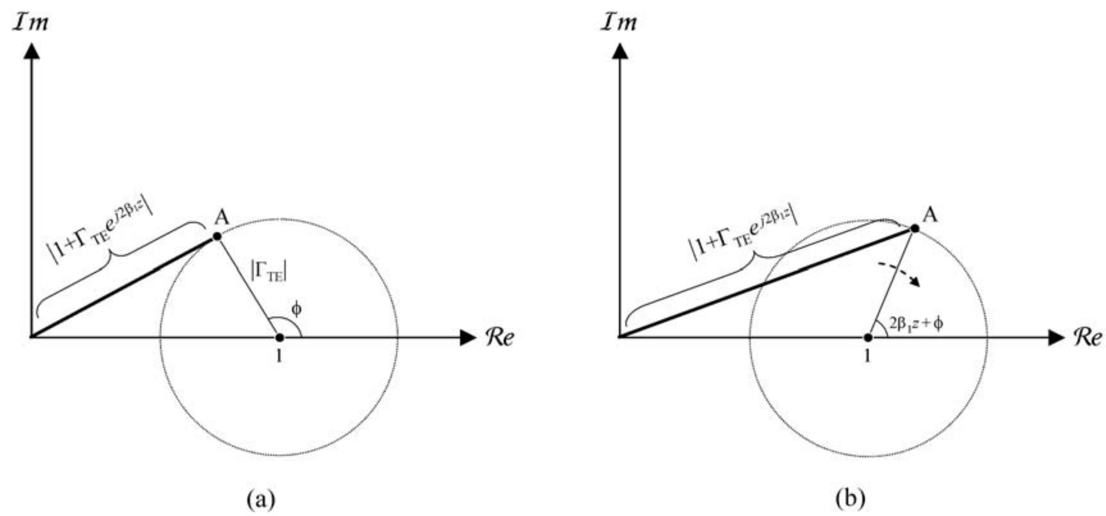


Figure 3.19: Visualization of the standing wave envelope using the crank diagram in the complex plane. The envelope is proportional to the length of an arm that extends from the origin to a rotating "crank" anchored at unity. (a) The crank at $z = 0$. (a) As the distance from the interface is increased (z becomes increasingly more negative), the crank rotates in a clockwise direction.

The largest SWR is achieved if $|\Gamma_{\text{TE}}| = 1$ which would occur, for example, if region 2 is a perfect conductor. In this case, the standing wave envelope will have a maximum value of $2E_i$ and a minimum value of zero. This results in a pure *standing wave*. For a general case in which $|\Gamma_{\text{TE}}| \neq 1$, the resulting pattern can be created by the superposition of a pure standing wave on a traveling wave.

Also from the crank diagram, we observe that the standing wave envelope is a periodic function. The period of the standing wave envelope, Λ, is the minimum distance in the z-direction required for the "crank" to return to the same position. This occurs when $2\beta_1\Lambda = 2\pi$, which yields

$$\Lambda = \frac{\pi}{\beta_1} = \frac{\pi}{n_1 k_o \cos\theta_1} = \frac{\lambda_o}{2n_1 \cos\theta_1} \tag{3.129}$$

since $k_o = 2\pi/\lambda_o$. As the incident angle increases, the period of the standing wave envelope increases. The standing wave envelope with the shortest period is achieved when the incident wave is at normal incidence ($\theta_1 = 0$). In this case, the period of the standing wave envelope is simply $\Lambda = \lambda_o/(2n_1)$.

Figure 3.20 illustrates the amplitude of the total electric field, $\overline{\mathcal{E}}_{\text{tot}}$, and the standing wave envelope for a wave that is normally incident on an interface. For this example, a reflection coefficient of $\Gamma = 0.5$ is assumed. Notice that as time is incremented, the electric field always stays within the standing wave envelope, which in this case has a SWR of $1.5/0.5 = 3$.

3.8 Plane Waves at Multiple Interfaces

Up to this point we have dealt with reflection and transmission from a single interface. However, it is easy to imagine a situation in which an electromagnetic wave will be incident on a region that consists of multiple, parallel interfaces. An example of this is sunlight incident on a puddle of water with a thin layer of oil floating on top. Another example of this is most eyeglass lenses, which are coated with thin layers of different materials to reduce reflections. If a medium consists of many parallel layers, we refer to it as a "multilayered" or "stratified" medium.

Previously, we developed a matrix analysis method to determine the reflection and transmission from a single interface. In this section, we extend the method to analyze reflection and transmission from a medium

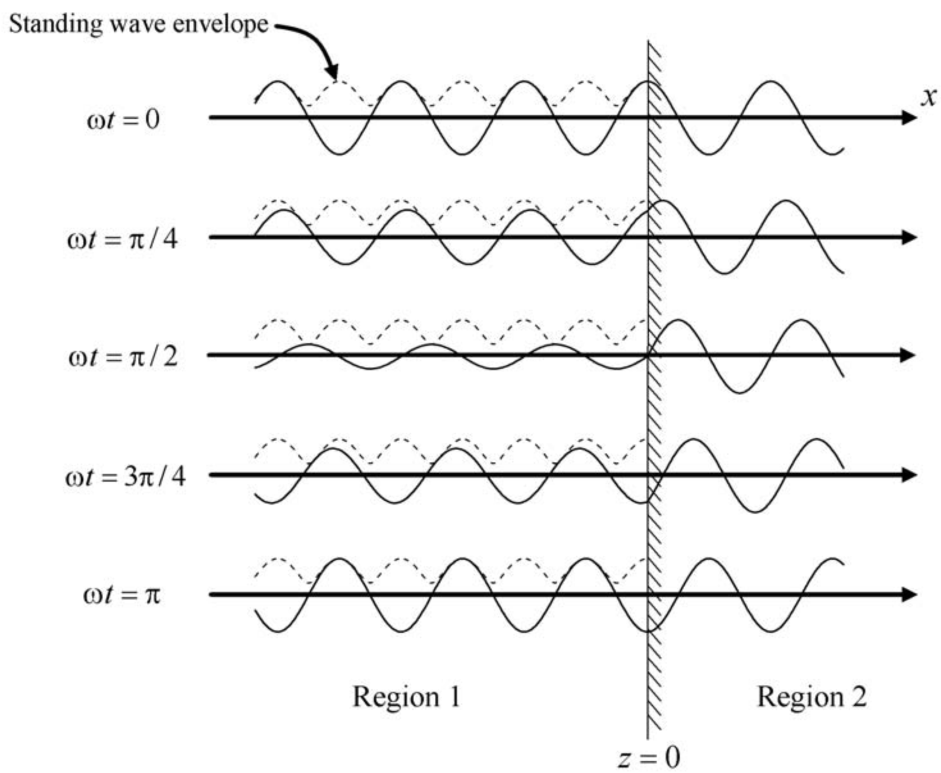


Figure 3.20: Electric field (solid lines) and standing wave envelope (dotted lines) for normal incidence onto an interface at $z = 0$. Notice that the electric field always stays within the envelope as the wave propagates. The reflection coefficient is $\Gamma = 0.5$.

with multiple layers. The approach is the following. In general, in each layer there will be a forward and backward travelling plane wave (the exception being the last layer in which there is only one wave, the transmitted wave). We have already derived relationships that relate the fields on one side of an interface to the other. If we can find a relationship to propagate two plane waves through a layer from one interface to the next over a certain distance, then we have all of the tools required to solve the problem. The solution method involves relating the plane waves across the first interface, propagating the plane waves through the next layer, relating the plane waves across the next interface, propagating through the next layer, and so forth. This process is illustrated in Figure 3.21. Mathematically, we will find that this process reduces to a series of matrix multiplications. Using this method, we can easily obtain the total reflection and transmission coefficients of the entire "stack" of layers. Note that in the previous section, we denoted the regions on either side of the interface as region 1 and region 2. For the case of multiple interfaces, we will number the layers 0, 1, 2, 3... and so on (the reason for numbering this way will become clear in the next chapter).

3.8.1 Relating Amplitudes Across One Interface

For the case of multiple interfaces, we will have layers 0, 1, 2, 3... and so on, and therefore there is an interface between regions 0 and 1, between regions 1 and 2, etc. By generalizing Equation 3.95, we can obtain a matrix expression that relates the amplitudes of the fields across an interface between any layer m

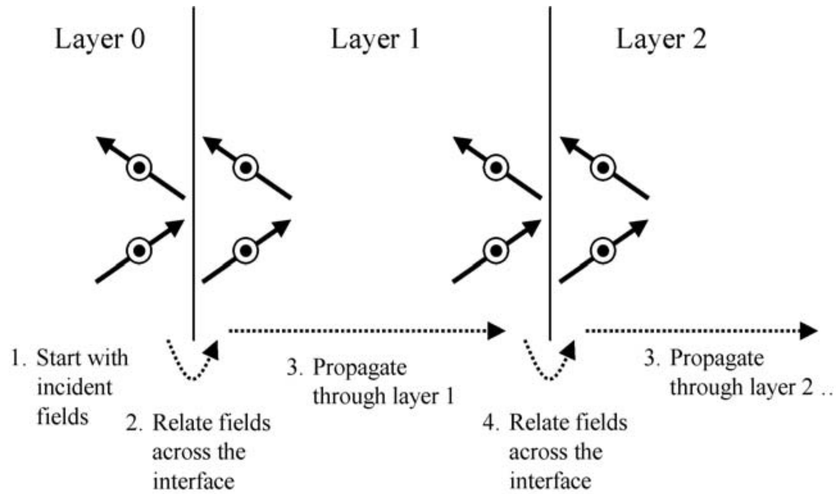


Figure 3.21: Process for solution of a multilayered problem. The solution method involves relating the plane waves across the first interface, propagating the plane waves through the next layer, relating the plane waves across the next interface, propagating through the next layer, and so forth.

and the next layer $m+1$:

$$\begin{aligned}\begin{pmatrix} A_m \\ B_m \end{pmatrix} &= \frac{1}{2}\begin{pmatrix} 1+\delta_{m(m+1)}\beta_{m+1}/\beta_m & 1-\delta_{m(m+1)}\beta_{m+1}/\beta_m \\ 1-\delta_{m(m+1)}\beta_{m+1}/\beta_m & 1+\delta_{m(m+1)}\beta_{m+1}/\beta_m \end{pmatrix}\begin{pmatrix} A_{m+1} \\ B_{m+1} \end{pmatrix} \\ &= \tilde{D}_{m(m+1)}\begin{pmatrix} A_{m+1} \\ B_{m+1} \end{pmatrix}\end{aligned} \tag{3.130}$$

where we have defined

$$\tilde{D}_{m(m+1)} = \frac{1}{2}\begin{pmatrix} 1+\delta_{m(m+1)}\beta_{m+1}/\beta_m & 1-\delta_{m(m+1)}\beta_{m+1}/\beta_m \\ 1-\delta_{m(m+1)}\beta_{m+1}/\beta_m & 1+\delta_{m(m+1)}\beta_{m+1}/\beta_m \end{pmatrix} \tag{3.131}$$

and recall that

$$\delta_{m(m+1)} = \begin{cases} 1 & \text{for TE-polarization} \\ \varepsilon_m/\varepsilon_{m+1} & \text{for TM-polarization} \end{cases} \tag{3.132}$$

Remember that A_m and B_m refer to the amplitudes of the forward and backward travelling *electric* fields if we are dealing with TE polarization, or the *magnetic* fields if we are dealing with TM polarization.

3.8.2 Relating Amplitudes Within a Layer

Within a particular layer m, we assume two plane waves: a forward travelling and a backward travelling plane wave. Because the waves are propagating over a certain distance in the z-direction, we no longer assume that the wave amplitudes, A_m and B_m, are constant. Instead, they are functions of z. Therefore, the total field in layer m has the general form

$$F_m = \left[A_m(z)e^{-j\beta_m z} + B_m(z)e^{+j\beta_m z}\right]e^{-j\kappa x} \tag{3.133}$$

where F_m represents the total *electric* field if we are dealing with the TE case or the total *magnetic* field for the TM case. Let us define the z-coordinate of the interface between layer $m-1$ and layer m as z_m, as shown in Figure 3.22. The total field in layer m at $z = z_m$ is then

$$F_m(z = z_m) = \left[A_m(z_m)e^{-j\beta_m z_m} + B_m(z_m)e^{+j\beta_m z_m}\right]e^{-j\kappa x} \tag{3.134}$$

If layer m has a thickness d_m, then the interface between layer m and the next layer $m+1$ occurs at $z = z_m + d_m$ (refer to Figure 3.22). The total field in layer m at $z = z_m + d_m$ is

$$F_m(z = z_m + d_m) = \left[A_m(z_m)e^{-j\beta_m(z_m+d_m)} + B_m(z_m)e^{+j\beta_m(z_m+d_m)}\right] e^{-j\kappa x}$$

$$= \left[A_m(z_m)e^{-j\beta_m d_m}e^{-j\beta_m z_m} + B_m(z_m)e^{+j\beta_m d_m}e^{+j\beta_m z_m}\right] e^{-j\kappa x}$$

$$= \left[A_m(z_m+d_m)e^{-j\beta_m z_m} + B_m(z_m+d_m)e^{+j\beta_m z_m}\right] e^{-j\kappa x} \tag{3.135}$$

Comparing Equations 3.134 and 3.135, we see that the field amplitudes within the layer are related by

$$A_m(z_m + d_m) = A_m(z_m)e^{-j\beta_m d_m} \tag{3.136}$$

$$B_m(z_m + d_m) = B_m(z_m)e^{+j\beta_m d_m} \tag{3.137}$$

In matrix form, this becomes

$$\begin{pmatrix} A_m(z_m+d_m) \\ B_m(z_m+d_m) \end{pmatrix} = \begin{pmatrix} e^{-j\beta_m d_m} & 0 \\ 0 & e^{+j\beta_m d_m} \end{pmatrix} \begin{pmatrix} A_m(z_m) \\ B_m(z_m) \end{pmatrix} = \tilde{P}_m\left(d_m\right) \begin{pmatrix} A_m(z_m) \\ B_m(z_m) \end{pmatrix} \tag{3.138}$$

where we have defined $\tilde{P}_m$, the ***propagation matrix***, as

$$\tilde{P}_m\left(d_m\right) = \begin{pmatrix} e^{-j\beta_m d_m} & 0 \\ 0 & e^{+j\beta_m d_m} \end{pmatrix} \tag{3.139}$$

This means that if we know the amplitudes of the fields just inside the layer, we can find the field amplitudes at the other side of the layer (after propagating a distance d_m) simply by multiplying by the propagation matrix $\tilde{P}_m(d_m)$. Later we will see that it is the *inverse* of the propagation matrix that is required for the multilayered formulation. Since $\tilde{P}_m$ is a diagonal matrix, the inverse is simply

$$\tilde{P}_m^{-1}\left(d_m\right) = \begin{pmatrix} e^{+j\beta_m d_m} & 0 \\ 0 & e^{-j\beta_m d_m} \end{pmatrix} \tag{3.140}$$

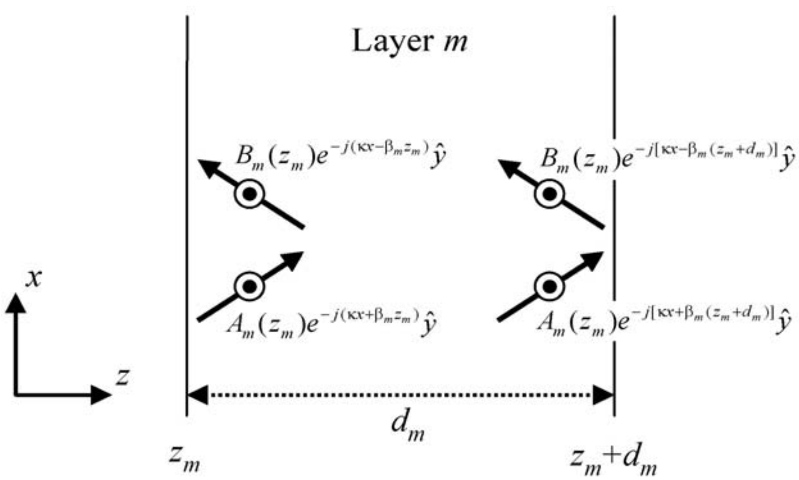


Figure 3.22: Propagation of the electromagnetic fields through layer m, which has a thickness d_m.

3.8.3 Relating Amplitudes Between Multiple Layers

We now examine the procedure for relating wave amplitudes between multiple layers in a stratified medium (a medium with many, parallel layers). The procedure is illustrated in Figure 3.23. To simplify the notation let $A_m(z_m) = A_m$, $B_m(z_m) = B_m$, and $\tilde{P}_m(d_m) = \tilde{P}_m$. We first number the layers in increasing order, with the incident layer as layer 0, then layer 1, layer 2, and so on until layer M. We assume that the amplitudes of the wave in the outermost layer (layer 0) are known, and are equal to A_0 and B_0 at the first interface. The amplitudes of the plane waves across the interface between layer 0 and layer 1 are then related by

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \tilde{D}_{01} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} \tag{3.141}$$

We next propagate the waves through layer 1 using the propagation matrix $\tilde{P}_1$ to obtain the wave amplitudes just before at the interface between regions 1 and 2:

$$\tilde{P}_1 \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \tilde{D}_{12} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.142}$$

Rearranging, we find

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \tilde{P}_1^{-1} \tilde{D}_{12} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.143}$$

Repeating this process for the interface at between regions 2 and 3 we obtain

$$\tilde{P}_2 \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \tilde{D}_{23} \begin{pmatrix} A_3 \\ B_3 \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \tilde{P}_2^{-1} \tilde{D}_{23} \begin{pmatrix} A_3 \\ B_3 \end{pmatrix} \tag{3.144}$$

This process is repeated until the last layer M, where we obtain

$$\tilde{P}_{M-1} \begin{pmatrix} A_{M-1} \\ B_{M-1} \end{pmatrix} = \tilde{D}_{(M-1)M} \begin{pmatrix} A_M \\ B_M \end{pmatrix} \quad \rightarrow \quad \begin{pmatrix} A_{M-1} \\ B_{M-1} \end{pmatrix} = \tilde{P}_{M-1}^{-1} \tilde{D}_{(M-1)M} \begin{pmatrix} A_M \\ B_M \end{pmatrix} \tag{3.145}$$

These matrix equations can be combined to relate the wave amplitudes between layer 1 and layer M:

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \tilde{D}_{01} \tilde{P}_1^{-1} \tilde{D}_{12} \tilde{P}_2^{-1} \tilde{D}_{23} ... \tilde{P}_{M-1}^{-1} \tilde{D}_{(M-1)M} \begin{pmatrix} A_M \\ B_M \end{pmatrix} \tag{3.146}$$

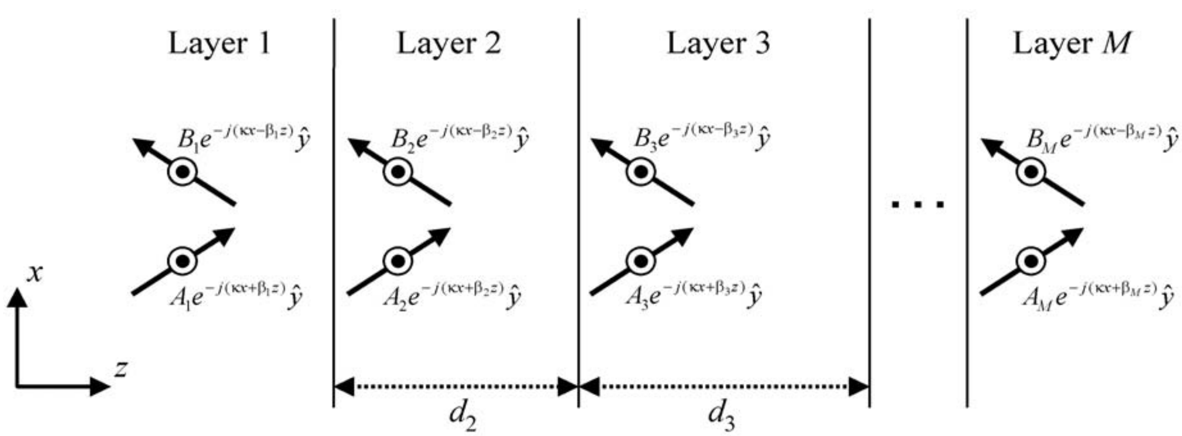


Figure 3.23: Illustration of the multilayer formulation.

After carrying out the matrix multiplication, the entire cascade of $\tilde{D}$ and $\tilde{P}$ matrices reduces to a single 2×2 matrix that relates the wave amplitudes in the first region (A_0, B_0) to the amplitudes in the last region (A_M, B_M). Similar to the single layer case, we then find the reflection and transmission coefficients from the entire stack of layers by considering only a transmitted wave in the last region (setting $B_M = 0$). Equation 3.146 then reduces to a matrix equation of the form:

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} A_M \\ 0 \end{pmatrix} \tag{3.147}$$

From this matrix equation we obtain $A_0 = m_{11} A_M$ and $B_0 = m_{21} A_M = (m_{21}/m_{11}) A_0$. Therefore, the reflection and transmission coefficients of the entire stack are

$$r = \frac{B_0}{A_0} = \frac{m_{21}}{m_{11}} \tag{3.148}$$

$$t = \frac{A_M}{A_0} = \frac{1}{m_{11}} \tag{3.149}$$

Remember that r and t are defined in terms of the electric fields for the TE case, and in terms of the magnetic fields for the TM case.

When implementing this method, the normal wave vector component, β_m, is required in each layer m to calculate the matrices $\tilde{D}$ and $\tilde{P}$. Assuming that we know the angle of incidence of the incident plane wave, θ_0, (where θ_0 is defined as the angle from the interface normal) it is straightforward to calculate κ and β_0, the x- and z-components of the wave vector in region 0:

$$\beta_0 = k_o n_0 \cos(\theta_0) \tag{3.150}$$

$$\kappa = k_o n_0 \sin(\theta_0) \tag{3.151}$$

where n_0 is the index of refraction in region 0. Since the tangential component of the wave vector, κ, must be the same in all layers, the normal component in each layer is found from

$$\beta_m = \sqrt{(k_o n_m)^2 - \kappa^2} \tag{3.152}$$

where n_m is the index of refraction of layer m, and $m = 1, 2, \ldots M$.

3.8.4 Example: Three Layer Case

The power of the matrix method for multilayered structures lies in the fact that computer routines can be quickly implemented to numerically calculate the reflection and transmission coefficients for a structure with an arbitrary number of layers. However, here we will consider the development of a general expression for the reflection and transmission coefficients of a three-layered structure using this method. This structure is illustrated in Figure 3.24. Our goal is to calculate the reflection coefficient from the stack in terms of the reflection and transmission coefficients r_{01}, t_{01}, r_{12}, t_{12}, and the thickness of the middle layer d_1.

Using the notation developed in the previous section, the relationships between the fields in layer 0 and layer 2 are expressed by the following matrix equation:

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \tilde{D}_{01} \tilde{P}_1^{-1} \tilde{D}_{12} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.153}$$

where (see Equation 3.100):

$$\tilde{D}_{01} = \frac{1}{t_{01}} \begin{pmatrix} 1 & r_{01} \\ r_{01} & 1 \end{pmatrix} \tag{3.154}$$

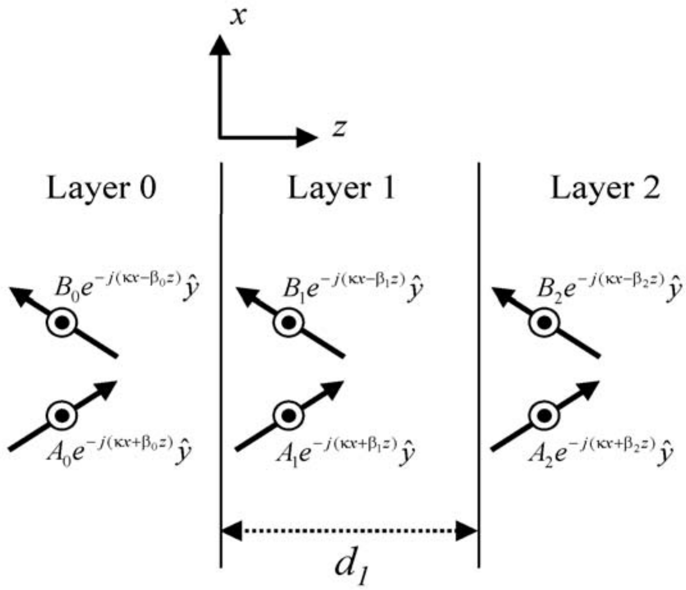


Figure 3.24: Illustration of the forward and backward travelling fields in a three-layered structure.

$$\tilde{P}_1^{-1} = \begin{pmatrix} e^{+j\beta_1 d_1} & 0 \\ 0 & e^{-j\beta_1 d_1} \end{pmatrix} \tag{3.155}$$

$$\tilde{D}_{12} = \frac{1}{t_{12}} \begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \tag{3.156}$$

For a practical calculation, we would carry out the matrix multiplication on a computer. However, for illustration purposes, we will explicitly perform the multiplication. Evaluating Equation 3.153 we obtain

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \frac{1}{t_{01}t_{12}} \begin{pmatrix} e^{j\beta_1 d_1} + r_{01}r_{12}e^{-j\beta_1 d_1} & r_{01}e^{-j\beta_1 d_1} + r_{12}e^{-j\beta_1 d_1} \\ r_{01}e^{j\beta_1 d_1} + r_{12}e^{-j\beta_1 d_1} & e^{-j\beta_1 d_1} + r_{01}r_{12}e^{-j\beta_1 d_1} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{3.157}$$

Since there is no backward travelling wave in layer 2, the reflection coefficient is

$$r = \frac{m_{21}}{m_{11}} = \frac{r_{01}e^{j\beta_1 d_1} + r_{12}e^{-j\beta_1 d_1}}{e^{j\beta_1 d_1} + r_{01}r_{12}e^{-j\beta_1 d_1}} = \frac{r_{01} + r_{12}e^{-j2\beta_1 d_1}}{1 + r_{01}r_{12}e^{-j2\beta_1 d_1}} \tag{3.158}$$

The transmission coefficient of the stack is

$$t = \frac{1}{m_{11}} = \frac{t_{01}t_{12}}{e^{j\beta_1 d_1} + r_{01}r_{12}e^{-j\beta_1 d_1}} = \frac{t_{01}t_{12}e^{-j\beta_1 d_1}}{1 + r_{01}r_{12}e^{-j2\beta_1 d_1}} \tag{3.159}$$

As an example, consider the case of a very thin layer of oil floating on a puddle of water. Because oil is less dense than water, a small amount of oil will float and can generate layers as thin as one micron or less. You may have observed this in a parking lot after a rain, and noticed different colors in the puddle. You may have also noticed that when you change the angle at which you are viewing the puddle, the apparent color changes. Why does this occur?

Figure 3.25 illustrates the case of a 0.5 μm thick layer of oil floating on water. For this example, we have assumed the indices of air, oil, and water are $n_0 = 1$, $n_1 = 1.7$, and $n_2 = 1.33$, respectively. Using Equation 3.158, we can calculate the reflection coefficient for various wavelengths and angles of incidence. Figure 3.26 illustrates the reflectivity for the TE case for three different incident angles (recall the reflectivity for the TE

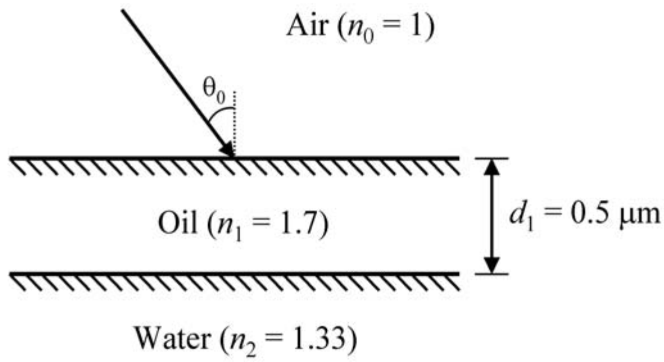


Figure 3.25: A thin layer of oil floating on top of water.

case is simply $R = |r|^2$). At normal incidence (0°), a peak in the reflectivity is observed at a wavelength of approximately 570 nm. In this case, the layer of oil on the puddle would appear greenish-yellow. If you are observing the puddle at an incident angle of 45°, a peak is observed at approximately 520 nm, and the puddle would appear bluish-green. If you are observing the puddle at an angle of 70°, the peak reflectivity occurs at about 470 nm. In this case, the puddle would appear blue.

3.9 Important Concepts

Upon the completion of this chapter, you should be able to:

- Recall the boundary conditions at an interface.
- Analyze plane waves at normal incidence onto a perfect conductor and a dielectric.
- Analyze plane waves at normal incidence onto a conductor, understand the generation of currents, and show how the analysis reduces to the case of a perfectly conducting interface.
- Explain why the tangential components of all plane waves at an interface must be equal.
- Use and understand the interpretation of "k-vector surfaces" at an interface.
- Define the following: plane of incidence, TE polarization and TM polarization.
- Understand why the solution is broken up into the solution of the TE and TM polarized components when analyzing the reflection and transmission of an arbitrary polarized wave, and why the solution is then the sum of the two polarizations.
- Derive the reflection and transmission coefficients for TE and TM polarized waves.
- Analyze the power flow at oblique incidence from first principles and show that power is conserved at an interface.
- Use k-vector surfaces to explain why and under which conditions total internal reflection occurs.
- Understand the behavior of fields and power flow when total internal reflection occurs.
- Understand Brewster's angle.
- Draw the reflection coefficient as a function of angle of incidence for a dielectric interface in which the incident medium has a lower or higher index of refraction than the second medium.

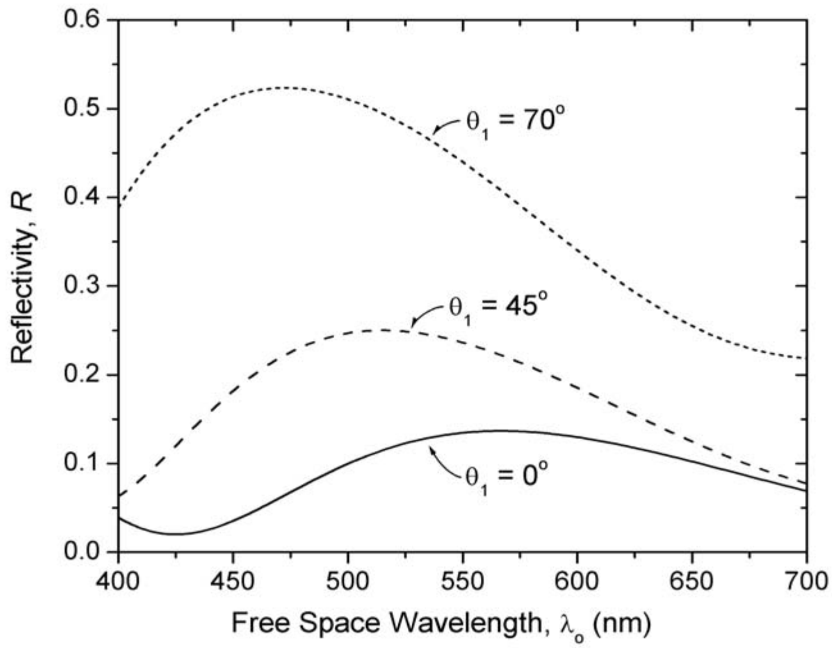


Figure 3.26: Reflectivity from a 0.5 μm thick layer of oil floating on water, TE polarization. (The indices of air, oil, and water are assumed to be $n_0 = 1$, $n_1 = 1.7$, and $n_2 = 1.33$, respectively). Notice how the wavelength of maximum reflectivity changes with incident angle.

- Mathematically describe the interference of two plane waves propagating at an angle with respect to each other.
- Formulate the reflection and transmission coefficients from a multilayered structure using the propagation matrix method.

Chapter 4

Waveguides

In the previous chapters we studied plane waves that propagate in an arbitrary direction in an infinite medium, or in a region with planar interfaces. Now we will look at structures specifically designed to guide electromagnetic waves, called "waveguides," that propagate a special class of electromagnetic fields called "modes." An electromagnetic mode is a *self-consistent solution of Maxwell's equations*, which means that once such a field solution exists in a waveguide, it will propagate without loss (unless, of course, the waveguide is made of an absorbing material). Waveguide modes propagate a limited amount of electromagnetic energy between two points in space in a secure manner without radiating energy. For this reason, electromagnetic modes are very important in high-speed point-to-point communication, such as in long-distance broadband internet connections. In this course we will focus on understanding the properties of modes, but will not concern ourselves with how these special field solutions are generated.

An important characteristic of waveguides is the ability to propagate a single mode at a specific carrier frequency with very low attenuation. The single mode must be able to exist over a large frequency band. In this manner, many signals at different frequencies can be propagated within the same waveguide over long distances with limited signal distortion. Another attractive characteristic of a waveguide is that a large amount of information can be transmitted in a very space-efficient manner. For example, a single optical cable may hold dozens of single mode optical fibers and within each fiber a very large number of signals can propagate (see Figure 4.1).

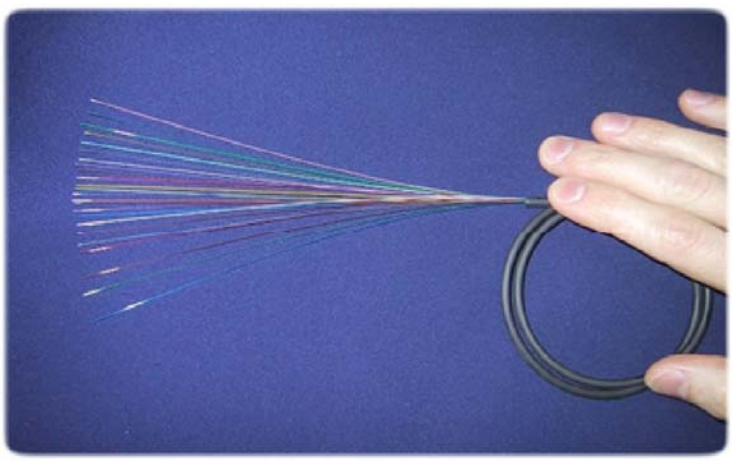


Figure 4.1: Optical cable, consisting of dozens of individual optical fibers. Each fiber strand is typically on the order of hundreds of microns in diameter [4].

Most waveguide structures fall into one of three categories: (1) transmission lines consisting of multiple conductors in which the dominant mode is a transverse electromagnetic wave, (2) closed conducting tubes, or (3) open-boundary dielectric structures, consisting of different dielectric materials. Types (1) and (2) are used mainly for transmitting high frequency electrical signals, while type (3) is used for optical signals. An example of a type (1) structure is a coaxial transmission line, which consists of a center metallic conductor, surrounded by a dielectric material and an outer conductor (see Figure 4.2). With this configuration, the electromagnetic mode is confined between the two conductors. A surprising fact is that a mode can also propagate if the center conductor of the coaxial transmission line is removed, which would yield a type (2) waveguide. The properties of such a mode in a type (2) waveguide is different than the transverse electromagnetic mode that propagates in the coaxial cable. An example of a type (2) waveguide is a rectangular metallic waveguide (see Figure 4.3) that is often used for guiding microwaves. What is even more surprising is that a mode can also propagate if both the inside and outside conducting regions are removed from a coaxial transmission line! This leaves only the dissimilar dielectric regions, which is a type (3) waveguide. An example of a type (3) waveguide is a glass optical fiber (see Figure 4.4). Glass fibers are used extensively for transmission of optical signals.

Figure 4.2: Coaxial cable. The outer diameter is typically on the order of 5 mm [5].

Figure 4.3: Rectangular metallic waveguides. The rectangular slot is typically on the order of 1 cm by 2 cm [6].

An electromagnetic mode has several similarities with a resonant LC circuit. In the case of an LC circuit, the resonance occurs in the time domain. This resonance is a "self-consistent" solution of the Kirchhoff's circuit equations. Once resonance begins, the LC circuit will continue to resonate at a specific frequency in the absence of losses. In a more general sense, it is possible for a ideal lossless LC oscillator to resonate at several distinct temporal frequencies. In this sense, the oscillator has "discretized" the time-frequency space.

A waveguide mode can be viewed as a "spatial-frequency" resonance. Unfortunately, the geometry of optical fibers and rectangular waveguides complicates the mathematical analysis, which can obscure the basic physical mechanisms. To gain an understanding of the fundamental physics behind waveguides, we will focus on the simplest waveguides consisting of planar dielectric or metallic regions. In a planar waveguide, the electromagnetic fields of guided modes can be written as a collection of plane waves. These plane waves can only propagate in specific directions within the waveguide, or alternatively stated, only specific wave vectors are allowed to propagate. In analogy to the resonant LC circuit, the electromagnetic mode is a resonance in *space*, and therefore it only occurs at specific, discrete wave vectors. In this way, it behaves like "spatial-frequency" resonance. There are also interesting analogies between guided electromagnetic waves and quantum theory. This is because of the analogy between the Maxwell wave equation and the Schrödinger equation, which is also a wave equation. A quantum course invariably starts off with studying a particle in a box (rectangular potential well), leading to resonances which explain why discrete quantum states exist in atoms. The allowed discrete states also occur in compound semiconductor structures, which in the two-dimensional case are described simply as stratified planar periodic media. This is analogous to

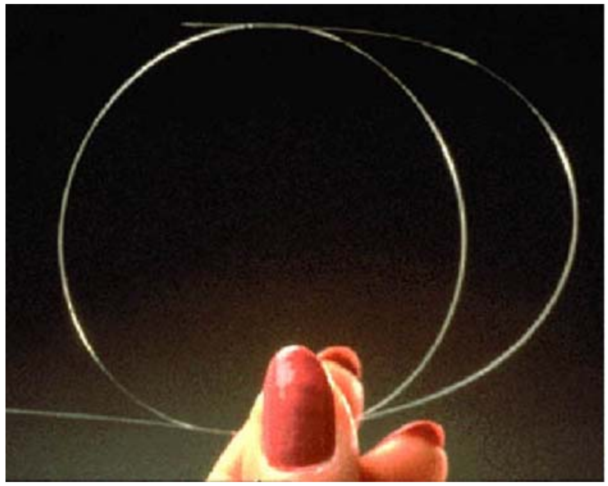


Figure 4.4: A single strand of optical fiber. The outer diameter of a fiber is typically on the order of hundreds of microns, including a protective plastic coating. The actual optical waveguide typically has a 125 micron outer diameter, and the region of the optical fiber that confines the light is on the order of 10 microns [7].

electromagnetic modes which, in contrast to plane waves that can propagate in any direction, i.e. *continuous* k-space propagation, have wave vectors that can only assume "discrete" values in k-space.

As stated previously, the study of planar waveguides is a good foundation for understanding modes of more complicated waveguide geometries such as rectangular guides and optical fibers. The simplest planar waveguide consists of one center planar region of finite width, surrounded by two "infinitely" thick regions. To treat different planar waveguides, we will derive a general condition for a planar waveguide mode that will then be simplified for two special cases. The first one is a planar waveguide in which the two outer regions are perfect conductors. This waveguide is called a parallel plate waveguide. The parallel plate waveguide propagates three types of modes: transverse electromagnetic (TEM) modes, transverse electric (TE) modes, and transverse magnetic (TM) modes. A detailed knowledge of the TE and TM modes in a parallel plate waveguide is important to understand TE and TM modes that exist in a more complicated rectangular waveguide. The second case that we will study consists of three dielectric regions. The outer regions have equal indices of refraction, while the inner layer has an index of refraction that is higher than the outer regions. Because the two outer regions have the same index of refraction, this waveguide is called a symmetric dielectric planar waveguide. This type of waveguide propagates only TE and TM modes (no TEM mode). A knowledge of planar dielectric waveguides is important for understanding optical fibers, since the modes in a dielectric waveguide resemble the modes in a optical fiber.

In the following sections, we will introduce the fundamental mechanisms that cause electromagnetic waves to be guided in a three-layered structure. We will then examine the specific cases of a metallic parallel plate waveguide and dielectric planar waveguide. Finally, we will extend the analysis to consider metallic waveguides that have a rectangular geometry. First, however, we must discuss an important change in the coordinate system used when dealing with waveguides. This is addressed in the next section.

4.1 The "Waveguide Coordinate System"

Up to this point, when discussing reflection and transmission of waves from an interface between two media we always chose the z-axis to be normal to the interface, as shown in Figure 4.5(a). This is because it is customary to chose the z-axis in the "general" direction of wave propagation, which has always been towards or away from the interface. Since we have defined the x-component of the wave vector as κ and the z-component as β, then κ was the tangential component, and β was the normal component. This meant that it was κ, the tangential component, that was the same in both regions.

However, we will see that when dealing with waveguides, the "general" direction of wave propagation is along the direction of the interface. Therefore, it is customary to choose the z-direction *parallel to the interface*! This means that now the x-direction is perpendicular to the interface, as shown in Figure 4.5(b)! As before, κ and β are defined as the x- and z-components of the wave vector. This means that now β is tangential to the interface and κ is normal to the interface, and it is β that must be equal in both regions.

Using the "waveguide coordinate system," the incident, reflected, and transmitted wave vectors shown in Figure 4.5(b) would be written as:

$$\bar{k}_i = \kappa_1 \hat{x} + \beta \hat{z} \tag{4.1}$$

$$\bar{k}_r = -\kappa_1 \hat{x} + \beta \hat{z} \tag{4.2}$$

$$\bar{k}_t = \kappa_2 \hat{x} + \beta \hat{z} \tag{4.3}$$

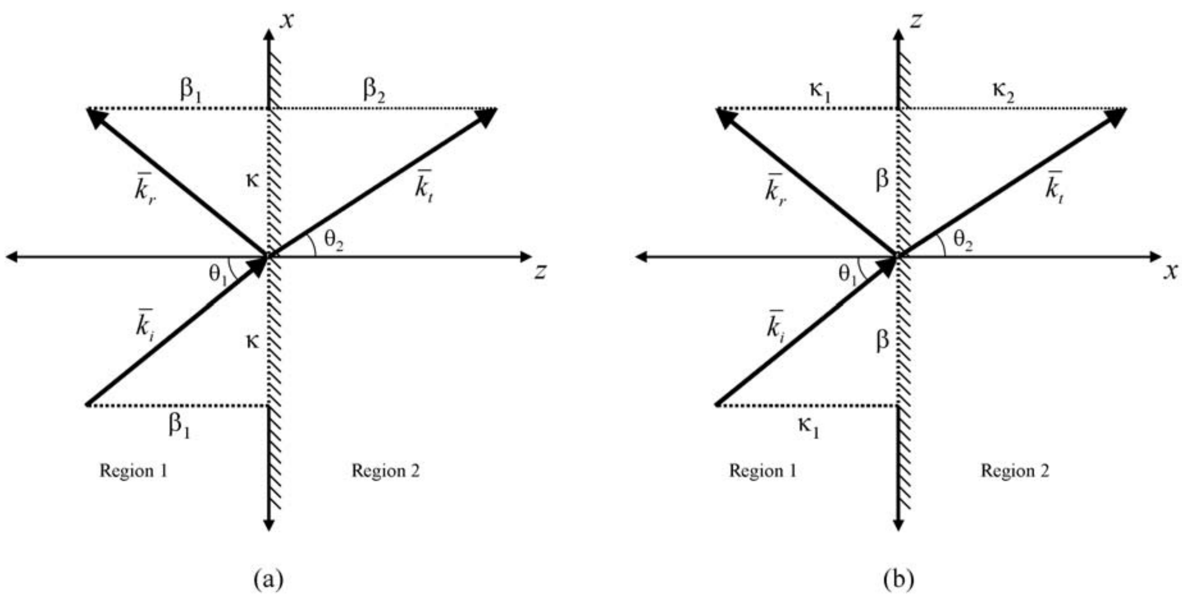


Figure 4.5: (a) Previously, the z-axis was chosen to be perpendicular to the interface. (b) In waveguiding applications, the general direction of propagation is along the interface, so the z-direction is defined parallel to the interface.

4.2 Waveguiding Condition for a Three-Layered Structure

Consider a structure consisting of three planar layers, which we will denote as region 0, region 1, and region 2. We are interested in electromagnetic modes propagating along the layers, and so we choose the waveguide coordinate system. Assume that region 0 extends to infinity in the negative x-direction, region 1 has a finite thickness *a,* and region 2 extends to infinity in the positive x-direction, as shown in Figure 4.6. As before, we assume a forward and a backward propagating plane wave in each region, with the amplitudes of the plane waves in the outermost regions related by (refer to Chapter 3):

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \tilde{D}_{01}\tilde{P}_1^{-1}\tilde{D}_{12}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{4.4}$$

where:

$$\tilde{D}_{01} = \frac{1}{t_{01}}\begin{pmatrix} 1 & r_{01} \\ r_{01} & 1 \end{pmatrix} \tag{4.5}$$

$$\tilde{P}_1^{-1} = \begin{pmatrix} e^{+j\kappa_1 a} & 0 \\ 0 & e^{-j\kappa_1 a} \end{pmatrix} \tag{4.6}$$

and

$$\tilde{D}_{12} = \frac{1}{t_{12}}\begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \tag{4.7}$$

Recall that $(k_o n_m)^2 = \kappa_m^2 + \beta^2$ in each layer m (the tangential wave vector component β is the same in

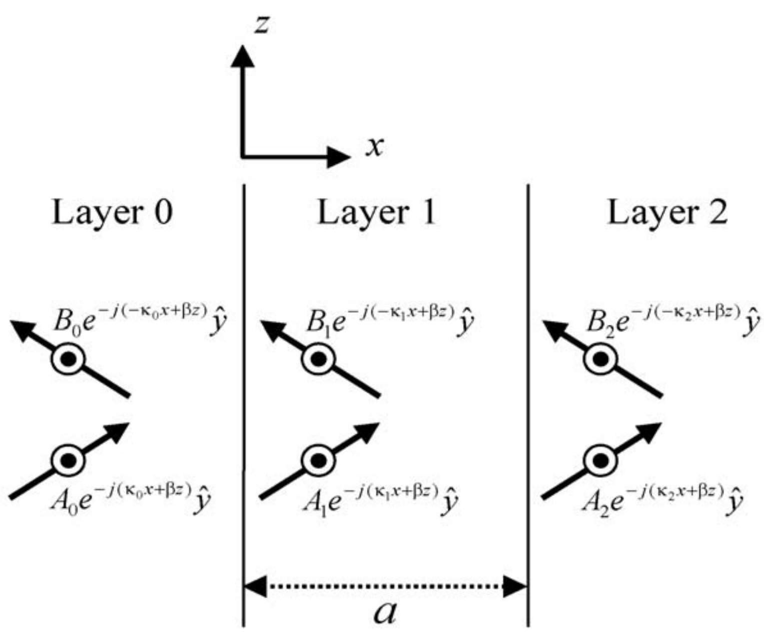


Figure 4.6: A three-layered planar waveguiding structure.

all three regions). Also remember that the amplitude A refers to a wave propagating in the forward $(+x)$ direction, and the amplitude B refers to a wave propagating in the backward $(-x)$ direction. We saw in the previous chapter that carrying out the matrix multiplication shown in Equation 4.4 yields

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \frac{1}{t_{01}t_{12}}\begin{pmatrix} e^{j\kappa_1 a} + r_{01}r_{12}e^{-j\kappa_1 a} & r_{01}e^{-j\kappa_1 a} + r_{12}e^{j\kappa_1 a} \\ r_{01}e^{j\kappa_1 a} + r_{12}e^{-j\kappa_1 a} & e^{-j\kappa_1 a} + r_{01}r_{12}e^{j\kappa_1 a} \end{pmatrix}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{4.8}$$

For brevity let us write this equation as

$$\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \tag{4.9}$$

To find the waveguide modes, we are interested in the *self-consistent solutions* of the matrix equation. A self-consistent solution is one in which waves exist in the center region without the presence of incident waves in the two outer regions. This means that we are interested in a solution where there is no wave incident in the +x-direction in layer 0 (no wave propagating in the forward direction, i.e $A_0 = 0$) and no wave incident in the -x-direction in layer 2 (no wave propagating in the backward direction, i.e. $B_2 = 0$). Substituting $A_0 = 0$ and $B_2 = 0$ into Equation 4.9 we find

$$0 = m_{11}A_2 \quad \text{and} \quad B_0 = m_{21}A_2 \tag{4.10}$$

The amplitudes A_2 and B_0 must be non-zero to prevent the trivial solution in which no fields are in the waveguide. Therefore, m_{11} must be zero. This leads to

$$m_{11} = e^{j\kappa_1 a} + r_{01}r_{12}e^{-j\kappa_1 a} = 0 \tag{4.11}$$

Multiplying both sides of the equation by $e^{-j\kappa_1 a}$ and using the fact that $r_{01} = -r_{10}$, Equation 4.11 is reduced to

$$r_{10}r_{12}e^{-j2a\kappa_1} = 1 \tag{4.12}$$

Therefore, an electromagnetic mode in a planar waveguide is the special electromagnetic field solution that satisfies Equation 4.12. In the case of the parallel plate waveguide, layers 0 and 2 are perfect conductors and therefore the reflection coefficients do not depend upon the angle of incidence. In this case a closed-form solution exists for Equation 4.12. In the case of a dielectric planar waveguide, the reflection coefficients are a function of the angle of incidence of the plane waves at the interfaces, which makes Equation 4.12 a nonlinear equation without a closed-form solution, and the resulting equation must be solved numerically.

4.3 Parallel Plate Waveguides

In this section, we analyze a waveguide that consists of a planar dielectric region (which may be air), bounded above and below by two conducting regions. To simplify the analysis, we will assume that the bounding regions are perfectly conducting. This approximation is valid in many practical situations, since the metals used for waveguides at microwave frequencies have very high conductivities.

We should point out here that when referring to waveguides that operate in the microwave regime such as parallel plate waveguides, it is customary to specify the permittivity, ε, of the dielectric region (or the relative permittivity, ε_r, where $\varepsilon_r = \varepsilon/\varepsilon_o$) and to think in terms of frequency. When discussing waveguides that are used in the optical regime, such as planar dielectric waveguides, it is customary to specify the index of refraction, n, of the dielectric regions, where $n = \sqrt{\varepsilon_r}$, and to think in terms of wavelength. Do not be confused by this, since it is easy to convert between these parameters and work in either the frequency or wavelength domain!

4.3.1 Propagation Constants of the Guided Modes

Consider a dielectric region bounded by two infinitely thick perfectly conducting planes separated by distance a, as shown in Figure 4.7. Assume that the region between the conductors (region 1) is a dielectric with a permittivity ε. Because regions 0 and 2 are perfectly conducting, a closed-form solution of the waveguiding condition (Equation 4.12) can be obtained. In Chapter 3, we found the reflection coefficients at the interface between a dielectric and a perfect conductor for the TE and TM cases. For the TE case, the reflection coefficient (defined in terms of the electric field) is $r = E_r/E_i = -1$. For the TM case, the reflection coefficient (defined in terms of the magnetic field) is $r = H_r/H_i = +1$. Referring to the waveguiding condition, we then note that the product $r_{10}r_{12}$ is *equal to one for both the TE and TM cases*! The waveguiding condition then reduces to

$$e^{-j2a\kappa_1} = \cos(2a\kappa_1) - j\sin(2a\kappa_1) = 1 \tag{4.13}$$

This equation is only satisfied when $2a\kappa_1 = 0, \pm 2\pi, \pm 4\pi, ...$ and so forth. This waveguiding condition can be summarized as

$$a\kappa_1 = m\pi \tag{4.14}$$

where m is defined as

$$m = \begin{cases} 1, 2, 3, 4, ... & \text{for TE modes} \\ 0, 1, 2, 3, ... & \text{for TM modes} \end{cases}$$

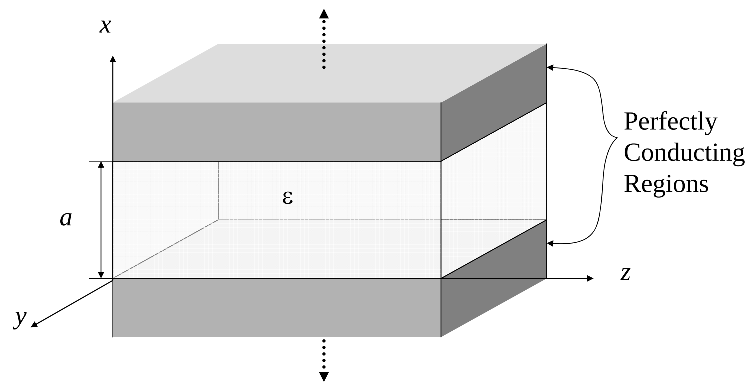


Figure 4.7: A waveguide consisting of a dielectric region surrounded by two infinitely thick perfectly conducting regions. It is assumed that the regions extend out to infinity in the x- and y-directions.

Each one of these discrete solutions corresponds to a different "mode" of the parallel plate waveguide (the reason that $m \neq 0$ for TE modes will be seen later.) This important relationship tells us that in a parallel plate waveguide, the allowed values of the transverse wave vector of a mode are determined only by separation of the plates! *The transverse wave vector of the parallel plate waveguide mode does not depend on the frequency/wavelength of operation, or the permittivity of the dielectric material!*

As the frequency of operation changes, the magnitude of the wave vector, k_1, also changes. But, the wave vector transverse to the plates, κ_1, must remain constant. Since $\kappa_1^2 + \beta^2 = k_1^2$ must be satisfied in the guiding region, this means that the wave vector component in the longitudinal direction, β, must change with frequency. This value of β associated with each mode at a particular frequency of operation is called the *propagation constant* of the mode. The propagation constant for a particular value of m is

$$\beta_m = \sqrt{k_1^2 - \kappa_1^2} = \sqrt{\omega^2 \mu_o \varepsilon - \left(\frac{m\pi}{a}\right)^2} = \sqrt{\frac{\omega^2 \varepsilon_r}{c^2} - \left(\frac{m\pi}{a}\right)^2} \tag{4.15}$$

The variation of the propagation constants with frequency for $m = 0$ to 4 are plotted in Figure 4.8 (recall that no $m = 0$ TE mode exists). Notice that only the $m = 0$ TM mode can propagate down to DC. In addition, notice that the $m = 0$ mode has a linear relationship between the frequency and propagation constant, while the other modes have nonlinear relationships. This will be important later when we discuss dispersion which leads to signal distortion in waveguides.

It should be pointed out that, although we initially assumed that the outer perfectly conducting regions are infinitely thick, this is not a necessary assumption. Because the electromagnetic fields are zero inside a perfect conductor, we may also assume that the perfectly conducting layers are infinitely thin (see Figure

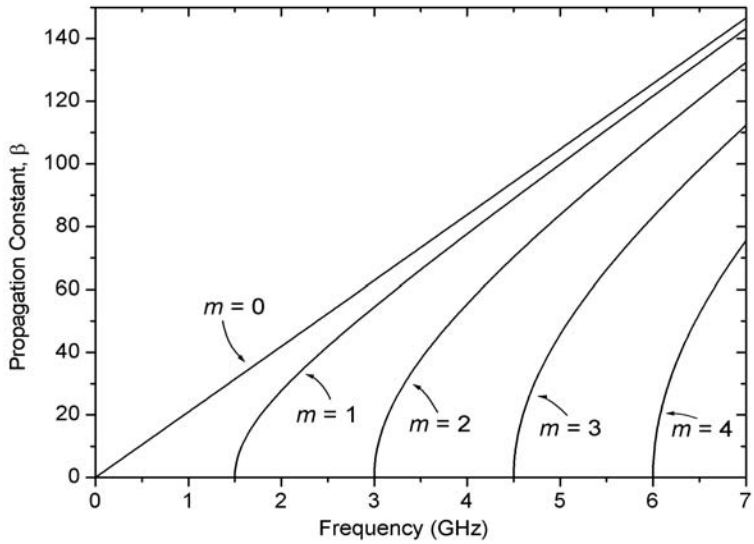


Figure 4.8: Propagation constants vs. frequency in a parallel plate waveguide. (Note that the $m = 0$ TE mode does not exist.) The $m = 0$ TM mode can propagate down to DC, whereas all of the other modes require a minimum frequency for propagation. (Waveguide parameters: $a = 10$ cm, $\varepsilon_r = 1$)

4.9). The previous derivation still holds and the waveguiding condition is the same. Of course, in real life there is no "infinitely thin perfect conductor," but this approximation is valid for very thin, highly conducting materials.

An alternative way to look at the waveguiding condition in a parallel plate waveguide is the following. Consider two plane waves polarized in the y-direction, i.e. TE waves, that are propagating down the parallel plate waveguide by "bouncing" off the top and bottom plates. Assume that the bounce angle is θ, as shown in Figure 4.10. At any point along the waveguide, the two waves can be written in phasor form as

$$\bar{E}_1 = A_1 e^{-j\beta z} e^{-j\kappa_1 x} \hat{y} \tag{4.16}$$

$$\bar{E}_2 = A_2 e^{-j\beta z} e^{+j\kappa_1 x} \hat{y} \tag{4.17}$$

where $\beta = k_1 \sin\theta$ and $\kappa_1 = k_1 \cos\theta$. The total electric field is then

$$\bar{E}_{\text{tot}} = \bar{E}_1 + \bar{E}_2 = A_1 e^{-j\beta z} e^{-j\kappa_1 x} \hat{y} + A_2 e^{-j\beta z} e^{+j\kappa_1 x} \hat{y} \tag{4.18}$$

We next impose the restriction that the wave amplitude must be zero at the lower perfectly conducting plate, located at $x = 0$. This will only be true for all values of κ_1 if $A_1 = -A_2 \equiv A$. Therefore, the total electric field can be written as

$$\bar{E}_{\text{tot}} = A\left(e^{-j\kappa_1 x} - e^{+j\kappa_1 x}\right) e^{-j\beta z} \hat{y} = -j2Ae^{-j\beta z} \sin(\kappa_1 x) \hat{y} \tag{4.19}$$

Notice that the standing wave generated by reflection from the lower plate has periodic locations along the x-axis where the field is always zero, given by $x = m\pi/\kappa_1$, where m is an integer. Therefore, the second perfectly conducting plate can be placed at any multiple of π/κ_1away from the first plate and the boundary conditions will be met on both plates. This can be expressed by the condition $a = m\pi/\kappa_1$ ($m = 1, 2, 3, ...$), which can be rearranged into the form given in Equation 4.14 above.

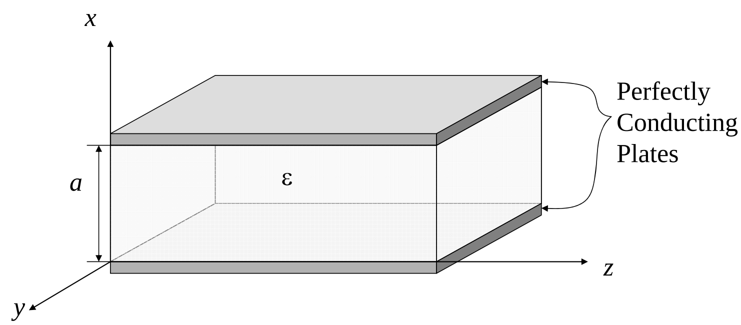


Figure 4.9: A practical parallel plate waveguide. The plates are thin, but the waveguiding condition still holds because the electromagnetic fields decay "infinitely fast" in a perfect conductor.

In summary, the field inside a parallel-plate waveguide can be viewed as a sum of two plane waves propagating at oblique angles. The relative phase of the two plane waves is such that the sum of the fields satisfy the boundary conditions at the interfaces. In the direction of propagation (i.e the z-direction, also called the *longitudinal direction*) the two plane waves have the same wave vector component, β. In the direction perpendicular to the interfaces (i.e the x-direction, also called the *transverse direction*) the magnitude of the wave vector components, κ_1, are equal but have opposite sign.

4.3.2 The "Discretization" of k-space

Now we can see why the wave vectors in a waveguide can only assume "discrete" values in k-space. The fields of plane waves propagating in infinite media and in "stratified" media can propagate in any direction, and the only restriction is that the wave vector must satisfy $\bar{k} \cdot \bar{k} = k^2$. Alternatively stated, there is no restriction on the direction of the wave vector, i.e. the k-space is continuous. In waveguides, however, the k-space is discretized. For example, in a parallel plate waveguide Equation 4.14 "discretizes" the k-space by restricting the wave vector components to the following values:

$$\kappa_1 \in \{0, \pi/a, 2\pi/a, 3\pi/a, ...\} \quad \text{for TM modes}$$
$$\kappa_1 \in \{\pi/a, 2\pi/a, 3\pi/a, ...\} \quad \text{for TE modes}$$

(The reason that $\kappa_1 \neq 0$ for TE modes will be seen in next section.) Figure 4.11 illustrates the discretization of k-space for TE modes due to the waveguiding condition of the parallel plate waveguide.

4.3.3 Cutoff Conditions for Parallel Plate Waveguides

Referring to Figure 4.8, we see that depending upon the frequency (or wavelength) of operation, a particular mode may or may not be able to propagate in a parallel plate waveguide. If a mode cannot propagate

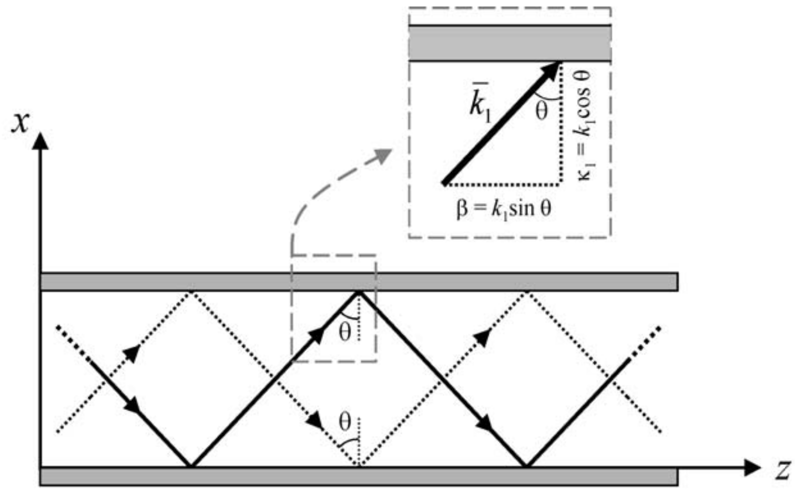


Figure 4.10: "Bouncing wave" interpretation of waveguiding by a parallel plate waveguide. The bounce angle of the waves is θ with respect to the normal of the perfectly conducting plates. The detail shows the decomposition of the wave vector k_1 into its normal and tangential components, κ_1 and β.

because the frequency of operation is too low (or the wavelength is too large) then the mode is referred to as being "*cut off.*"

Let us look at what happens to the TE and TM $m = 1$ modes as the frequency is decreased. From Equation 4.14 we see that the transverse wave vector must be equal to $\kappa_1 = \pi/a$ for the $m = 1$ mode. If the frequency of operation is low enough such that $k_1 < \pi/a$, then it is impossible for the transverse wave vector to equal π/a since the relationship $k_1^2 = \kappa_1^2 + \beta^2$ must be satisfied in the guiding region. For this reason, the $m = 1$ mode is cut off. The $m = 1$ *cutoff frequency* is defined as the *lowest* frequency at which the $m = 1$ mode can propagate without loss. Therefore, the cutoff frequency of a mode occurs when $\beta = 0$ and $k_1 = \kappa_1$. For the $m = 1$ mode, the cutoff condition is then

$$k_1 = \frac{2\pi f_{c,1}\sqrt{\varepsilon_r}}{c} = \frac{\pi}{a} \tag{4.20}$$

which, when rearranged, yields

$$f_{c,1} = \frac{c}{2a\sqrt{\varepsilon_r}} \tag{4.21}$$

where $f_{c,1}$ is the cutoff frequency of the $m = 1$ mode. This can be rewritten in terms of the wavelength as $\lambda_{c,1} = 2a\sqrt{\varepsilon_r}$, where $\lambda_{c,1}$ is the *cutoff wavelength in free space* (not the wavelength in the dielectric material, which would be shorter by a factor of $1/\sqrt{\varepsilon_r}$).

This condition can be generalized to all modes. The cutoff condition for the m^{th} mode in a parallel plate waveguide is

$$f_{c,m} = \frac{mc}{2a\sqrt{\varepsilon_r}} \qquad \text{or} \qquad \lambda_{c,m} = \frac{2a\sqrt{\varepsilon_r}}{m} \tag{4.22}$$

where $f_{c,m}$ is the *cutoff frequency of the* m^{th} *mode*, and $\lambda_{c,m}$ is the *cutoff wavelength in free space (not the wavelength inside the guide).* If we return to the view of the fields inside the waveguide as the sum of two "bouncing" plane waves, at cutoff these waves bounce back and forth at normal incidence between the plates since $\beta = 0$. In other words, the mode does not propagate in the z-direction. As the frequency is increased, the plane waves begin propagating in the z-direction.

The cutoff frequency/wavelength of a mode is an important parameter of a waveguide. If a mode is below the cutoff frequency, the mode cannot propagate and will be attenuated. Figures 4.12(a) and (b) show the

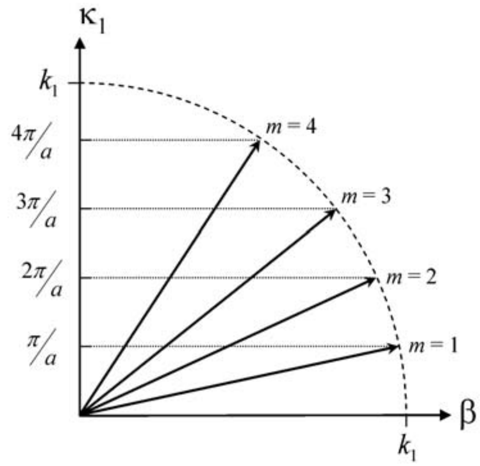


Figure 4.11: The discretization of k-space for TE modes in a parallel plate waveguide (note there is no $m = 0$ TE mode). The vectors shown represent the wave vectors for the $m = 1$, 2, 3, and 4 modes. The wave vector component in the transverse direction is fixed by the waveguide geometry, while the magnitude of the wave vector varies with frequency as $k_1 = \omega\sqrt{\mu_o \varepsilon}$.

number of propagating modes in a parallel plate waveguide versus frequency for TE modes and TM modes, respectively. Note that, at a specific frequency, there will always be one more TM mode than TE mode, since no $m = 0$ TE mode exists.

Using the expressions for the cutoff frequency and wavelength of each mode (Equation 4.22), we can rewrite the propagation constant of the m^{th} mode as (see Appendix D):

$$\beta_m = k_1\sqrt{1 - f_{c,m}^2/f^2} = k_1\sqrt{1 - \lambda^2/\lambda_{c,m}^2} \tag{4.23}$$

Remember that the cutoff wavelength in Equations 4.22 and 4.23 is the wavelength in free space, not the wavelength inside the waveguide!

Often, when speaking of metallic waveguides, the cutoff wavelength is given in terms of the wavelength *inside of the waveguide*. Since the wavelength inside of the waveguide is related to the wavelength outside by $\lambda_{\text{in}} = \lambda_{\text{out}}/\sqrt{\varepsilon_r}$, this means that the cutoff wavelength specified in terms of the wavelength inside of the waveguide does not depend upon the material in the waveguide! This can lead to confusion when specifying the cutoff wavelength of a parallel plate waveguide. This point is illustrated in the following example.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 4.1:

Consider a parallel plate waveguide filled with a nonmagnetic medium with a permittivity $\varepsilon = 9\varepsilon_o$. The spacing between the plates is 10 cm. What are the cutoff frequency and cutoff wavelength of the $m = 2$ TE mode? What are the cutoff frequency and cutoff wavelength if the permittivity of the filling material is increased to $\varepsilon = 16\varepsilon_o$? What is the propagation constant of the TE_2 mode at 1.5 GHz in each case?

Solution:

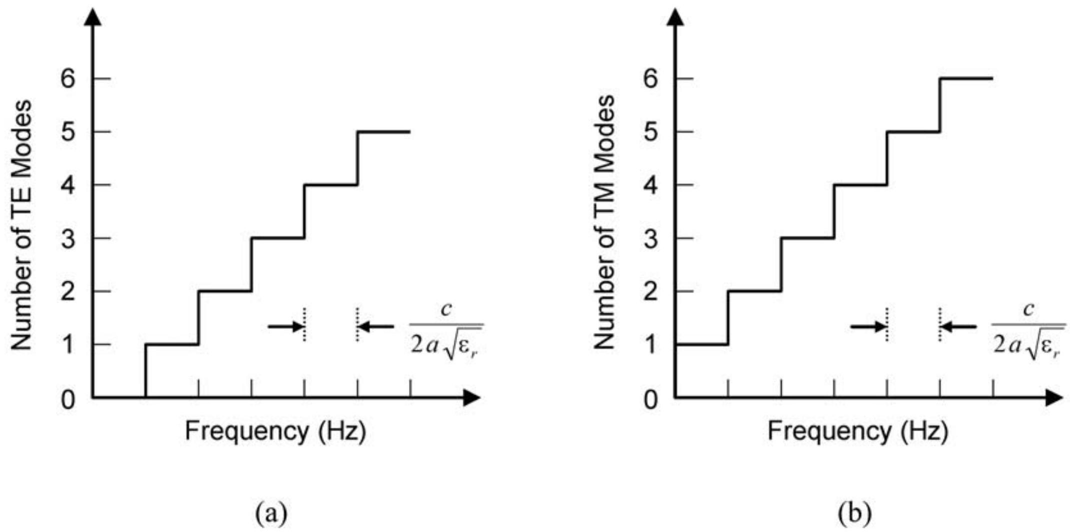


Figure 4.12: (a) The number of TE modes as a function of frequency in a parallel plate waveguide. The lowest-order mode is the $m = 1$ mode which has a cutoff frequency of $f_{c,1} = c/(2a\sqrt{\varepsilon_r})$. (b) The number of TM modes as a function of frequency in a parallel plate waveguide. The lowest-order mode is the $m = 0$ mode which can propagate down to 0 Hz.

The cutoff frequencies for the two different filling materials are given by Equation 4.22:

$$f_{c,2} = \frac{mc}{2a\sqrt{\varepsilon_r}} = \frac{2\left(3 \times 10^8\right)}{2\,(0.10)\,\sqrt{9}} = 1\,\text{GHz} \qquad (\text{for}\,\varepsilon = 9\varepsilon_o)$$

$$f_{c,2} = \frac{mc}{2a\sqrt{\varepsilon_r}} = \frac{2\left(3 \times 10^8\right)}{2\,(0.10)\,\sqrt{16}} = 0.75\,\text{GHz} \qquad (\text{for}\,\varepsilon = 16\varepsilon_o)$$

The cutoff wavelengths are

$$\lambda_{c,2} = \frac{2a\sqrt{\varepsilon_r}}{m} = \frac{2\,(0.10)\,\sqrt{9}}{2} = 30\,\text{cm} \qquad (\text{for}\,\varepsilon = 9\varepsilon_o)$$

$$\lambda_{c,2} = \frac{2a\sqrt{\varepsilon_r}}{m} = \frac{2\,(0.10)\,\sqrt{16}}{2} = 40\,\text{cm} \qquad (\text{for}\,\varepsilon = 16\varepsilon_o)$$

These are the cutoff wavelengths outside of the waveguide, i.e. in free space. Inside the waveguide, the wavelengths are decreased by a factor of $1/\sqrt{\varepsilon_r}$:

$$\lambda_{c,2}\,(\text{inside}) = \frac{\lambda_{c2}}{\sqrt{\varepsilon_r}} = \frac{30}{\sqrt{9}} = 10\,\text{cm} \qquad (\text{for}\,\varepsilon = 9\varepsilon_o)$$

$$\lambda_{c,2}\,(\text{inside}) = \frac{\lambda_{c2}}{\sqrt{\varepsilon_r}} = \frac{40}{\sqrt{16}} = 10\,\text{cm} \qquad (\text{for}\,\varepsilon = 16\varepsilon_o)$$

Notice that the cutoff wavelength inside of the waveguide does not depend upon the properties of the filling material! The propagation constant of the TE_2 mode at 1.5 GHz is:

$$\beta_2 = k_1\sqrt{1 - f_{c,2}^2/f^2} = \frac{2\pi \times 1.5 \times 10^9\sqrt{9}}{3 \times 10^8}\sqrt{1 - 1^2/1.5^2}$$

$$= 70.2 \quad \text{rad/m} \qquad (\text{for}\, \varepsilon = 9\varepsilon_o)$$

$$\beta_2 = k_1\sqrt{1 - f_{c,2}^2/f^2} = \frac{2\pi \times 1.5 \times 10^9\sqrt{16}}{3 \times 10^8}\sqrt{1 - 0.75^2/1.5^2}$$

$$= 109 \quad \text{rad/m} \qquad (\text{for}\, \varepsilon = 16\varepsilon_o)$$

In both cases, however, the transverse component of the wave vector must be

$$\kappa_1 = \frac{m\pi}{a} = \frac{2\pi}{0.10} = 62.8 \quad \text{rad/m}$$

Thinking in terms of the "bouncing wave" model, we observe that the "bounce angle" is steeper in the case in which $\varepsilon = 9\varepsilon_o$, since in that case the mode is closer to cutoff (refer to Figure 4.13).

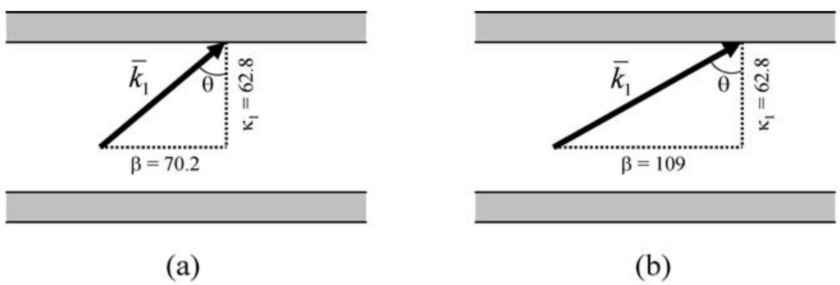


Figure 4.13: Wavevectors for the two different filling materials of Example 4.1: (a) $\varepsilon = 9\varepsilon_o$, (b) $\varepsilon = 16\varepsilon_o$.

⋘⋘⋘⋘⋙⋙⋙⋙

4.3.4 Modal Behavior Below Cutoff

It was previously stated that a mode cannot propagate without loss below the cutoff frequency of that particular mode. Below the cutoff frequency, the fields inside the waveguide have some similarities to fields in the case of total internal reflection (TIR). In TIR, the tangential wave vector component of the incident wave in the first medium is so large that in the second, low-index region the wave vector component perpendicular to the interface is forced to be imaginary. In the low-index region, the fields decay away from the interface, since the wave vector component perpendicular to the interface is imaginary, while the wave vector component parallel to the interface is real.

In a parallel plate waveguide, the transverse wave vector component of the mode remains equal to $\kappa_1 = m\pi/a$. Below cutoff, this means that $\kappa_1 > k_1$, so that the propagation constant in the z-direction must become imaginary. This means that the field *attenuates in the z-direction.* The further the frequency is below cutoff, the faster the mode attenuates.

If a mode is below the cutoff frequency (i.e. if $f < f_{c,m}$ or equivalently if $\lambda > \lambda_{c,m}$) then the propagation constant is imaginary and the mode is attenuated. The attenuation constant of the mode is

$$\alpha_m = \sqrt{\kappa^2 - k_1^2} = k_1\sqrt{f_{c,m}^2/f^2 - 1} = k_1\sqrt{\lambda^2/\lambda_{c,m}^2 - 1} \tag{4.24}$$

where we have defined $\alpha_m = j\beta_m$ as the attenuation constant of the m^{th} mode, and $k_1 = \omega\sqrt{\mu_o\varepsilon} = \omega\sqrt{\varepsilon_r}/c$.

⋘⋘⋘⋘⋙⋙⋙⋙

EXAMPLE 4.2:

Consider a parallel plate waveguide from Example 4.1 filled with a nonmagnetic medium with a permittivity $\varepsilon = 9\varepsilon_o$. The waveguide has a spacing between the plates of 10 cm. What is the attenuation in dB/cm of the $m = 2$ mode if we are operating at 200 MHz?

Solution:

In the previous example, we found that the cutoff frequency of the $m = 2$ mode is 1 GHz. Therefore, the attenuation constant is

$$\alpha_2 = k_1\sqrt{f_{c,m}^2/f^2 - 1} = (\omega\sqrt{\varepsilon_r}/c)\sqrt{f_{c,m}^2/f^2 - 1}$$

$$= \left(2 \times \pi \times 200 \times 10^6 \times \sqrt{9}/3 \times 10^8\right)\sqrt{1^2/0.2^2 - 1} = 61.6 \quad \text{Np}/\text{m}$$

Since the mode will attenuate as $E_o e^{-\alpha z}$, we can find the field amplitude at 0 cm and at 1 cm:

$$E(0\,\text{cm}) = E_o \quad \text{and} \quad E(1\,\text{cm}) = E_o e^{-(61.6)(0.01)} = 0.54 E_o$$

The attenuation in dB/cm is then

$$20\log\left(\frac{E_o}{0.54E_o}\right) = 5.35 \quad \text{dB}/\text{cm}$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

4.3.5 Wavelength and Phase Velocity in the Waveguide

In addition to the operating frequency ω and operating wavelength λ, we are also interested the wavelength and phase velocity in the direction of propagation (i.e. the z-direction). Using the expressions for the cutoff frequency and wavelength of each mode (Equation 4.22), we can define λ_z as the wavelength in the z-direction, or "guide wavelength" for the m^{th} mode (see Appendix D):

$$\lambda_{z,m} = \frac{2\pi}{\beta_m} = \frac{\lambda/\sqrt{\varepsilon_r}}{\sqrt{1 - f_{c,m}^2/f^2}} = \frac{\lambda/\sqrt{\varepsilon_r}}{\sqrt{1 - \lambda^2/\lambda_{c,m}^2}} \tag{4.25}$$

The phase velocity of the m^{th} mode is (see Appendix D):

$$v_{p,m} = \frac{\omega}{\beta_m} = \frac{c/\sqrt{\varepsilon_r}}{\sqrt{1 - f_{c,m}^2/f^2}} = \frac{c/\sqrt{\varepsilon_r}}{\sqrt{1 - \lambda^2/\lambda_{c,m}^2}} \tag{4.26}$$

4.3.6 Electromagnetic Fields: TE Modes

We now derive the electromagnetic field solutions inside the parallel plate waveguide. Using the "bouncing wave model," we have already derived the electric field for TE modes (Equation 4.19). Grouping all of the constant terms into the arbitrary constant E_o, the electric field can be rewritten as

$$\bar{E} = E_o \sin(\kappa_1 x) e^{-j\beta z}\hat{y} = E_o \sin\left(\frac{m\pi}{a}x\right) e^{-j\beta z}\hat{y} \tag{4.27}$$

where we have substituted in the waveguiding condition $\kappa_1 = m\pi/a$. For the TE modes, there is a problem if $m = 0$. If we choose $m = 0$, then the electric field is $\bar{E} = 0$. This is because the value of m determines the number of maxima/minima of the mode in the transverse direction. The $m = 0$ mode has zero maxima/minima, i.e. it is a constant in the transverse direction. But, the tangential electric field at the plates must be zero, so that a TE mode with a constant amplitude in the transverse direction cannot exist. This means that in a parallel plate waveguide the $m = 0$ TE mode cannot exist!

The magnetic field solution is easily obtained from the electric field using Faraday's law in time-harmonic form. (Note that we cannot use the expression $\bar{H} = \hat{k} \times \bar{E}/\eta$ to find the magnetic field, since the total electric field consists of *two* plane waves propagating in different directions!):

$$\bar{H} = \frac{j}{\omega\mu_o}\nabla \times \bar{E} = \frac{j}{\omega\mu_o}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & E_y & 0 \end{vmatrix} = \frac{j}{\omega\mu_o}\left(-\frac{\partial E_y}{\partial z}\hat{x} + \frac{\partial E_y}{\partial x}\hat{z}\right)$$

$$= \frac{E_o}{\omega\mu_o}\left\{-\beta\sin(\kappa_1 x)\hat{x} + j\kappa_1\cos(\kappa_1 x)\hat{z}\right\}e^{-j\beta z}$$

$$= \frac{E_o}{\omega\mu_o}\left\{-\beta\sin\left(\frac{m\pi}{a}x\right)\hat{x} + j\frac{m\pi}{a}\cos\left(\frac{m\pi}{a}x\right)\hat{z}\right\}e^{-j\beta z} \tag{4.28}$$

where $m \geq 1$ for TE modes. The electric fields of the three lowest-order TE modes are illustrated in Figure 4.14.

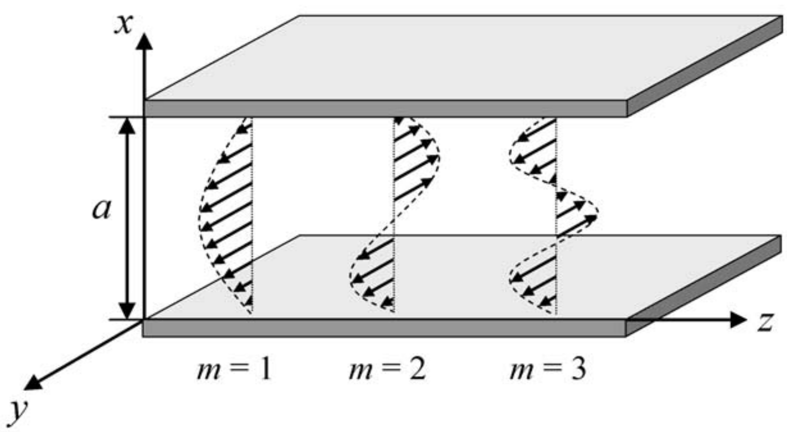


Figure 4.14: The electric fields of the three lowest-order TE modes in a parallel plate waveguide (note that the $m = 0$ mode cannot exist for the TE case). The arrows indicate the direction of the electric field vector (y-direction).

4.3.7 Electromagnetic Fields: TM Modes

For the TM modes, we again assume the "bouncing wave" model, except now the magnetic field is perpendicular to the direction of propagation. Therefore, we construct the magnetic field from the sum of two waves:

$$\bar{H}_{\text{tot}} = \bar{H}_1 + \bar{H}_2 = A_1 e^{-j\beta z}e^{-j\kappa_1 x}\hat{y} + A_2 e^{-j\beta z}e^{+j\kappa_1 x}\hat{y} \tag{4.29}$$

We cannot directly use the tangential magnetic fields to solve for the relationship between A_1 and A_2 because there are unknown surface currents on the perfect conductors. This means that the tangential magnetic field may not be zero at the surface of the perfect conductors. The problem is solved by first finding the tangential electric field at the interface and exploiting the fact that the total tangential electric field is zero at a perfect conductor.

Because the magnetic field vector is in the y-direction, the electric field will have components in the x- and z-directions. Since we are interested only in the tangential electric field, we use Ampere's law in time-harmonic form ($\nabla \times \bar{H} = j\omega\varepsilon\bar{E}$) to find the z-component of the electric field:

$$E_z = -\frac{j}{\omega\varepsilon}\frac{\partial H_y}{\partial x} = -\frac{\kappa_1}{\omega\varepsilon}\left(A_1 e^{-j\kappa_1 x} - A_2 e^{+j\kappa_1 x}\right)e^{-j\beta z} \tag{4.30}$$

Setting the tangential electric field to zero at $x = 0$, it follows that $A_1 = A_2$. Therefore the magnetic field can be written as

$$\bar{H} = A_1(e^{-j\kappa_1 x} + e^{+j\kappa_1 x})e^{-j\beta z}\hat{y} = H_o \cos(\kappa_1 x)e^{-j\beta z}\hat{y}$$

$$= H_o \cos\left(\frac{m\pi}{a}x\right)e^{-j\beta z}\hat{y} \tag{4.31}$$

Contrary to the TE mode, m can be zero for a TM mode. Substituting $m = 0$ into Equation 4.31 we find the magnetic field solution for the $m = 0$ TM mode is $\bar{H} = H_o e^{-j\beta z}\hat{y}$. This does not pose a problem since at a perfectly conducting interface the tangential magnetic field does not have to be equal to zero. Recall that a nonzero tangential magnetic field gives rise to an infinitely thin surface current flowing at the interface of the perfect conductor. The magnetic fields of the three lowest-order TM modes are shown in Figure 4.15.

Again using Ampere's law in time-harmonic form, we find that the electric field is

$$\bar{E} = \frac{-j}{\omega\varepsilon}\nabla \times \bar{H} = \frac{-j}{\omega\varepsilon}\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & H_y & 0 \end{vmatrix} = \frac{j}{\omega\varepsilon}\left(\frac{\partial H_y}{\partial z}\hat{x} - \frac{\partial H_y}{\partial x}\hat{z}\right)$$

$$= \frac{H_o}{\omega\varepsilon}\left\{\beta\cos(\kappa_1 x)\hat{x} + j\kappa_1\sin(\kappa_1 x)\hat{z}\right\}e^{-j\beta z}$$

$$= \frac{H_o}{\omega\varepsilon}\left\{\beta\cos\left(\frac{m\pi}{a}x\right)\hat{x} + j\frac{m\pi}{a}\sin\left(\frac{m\pi}{a}x\right)\hat{z}\right\}e^{-j\beta z} \tag{4.32}$$

where $m \geq 0$ for TM modes.

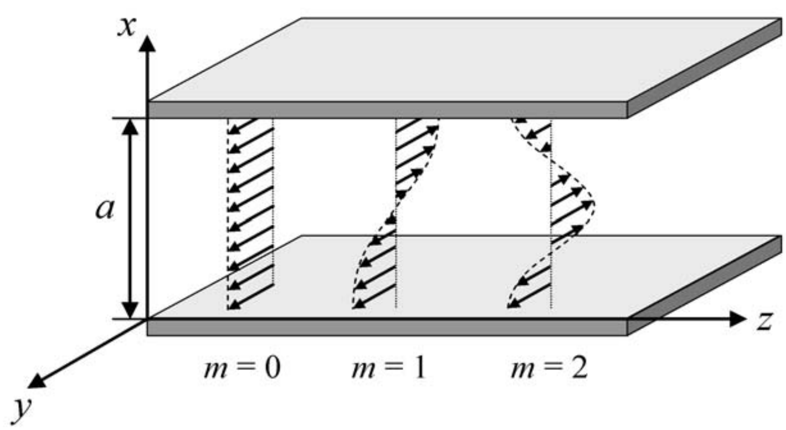


Figure 4.15: The magnetic fields of the three lowest-order TM modes in a parallel plate waveguide. Note the presence of an $m = 0$ TM mode, since the magnetic field can be constant without violating the boundary conditions at the plates. The arrows indicate the direction of the magnetic field vector (y-direction).

4.3.8 The TEM Mode

The lowest-order ($m = 0$) TM mode is a very special and important electromagnetic mode of the parallel plate waveguide. This mode has a zero-frequency (DC) cutoff. From Equation 4.14 we see that $\kappa_1 = 0$ for this mode. The propagation constant is then (see Equation 4.15)

$$\beta_o = k_1 = \omega\sqrt{\mu_o \varepsilon} \tag{4.33}$$

In other words, the propagation constant is that of a plane wave propagating through a material with a permittivity ε! The electric field for this mode is found by substituting $m = 0$ into Equation 4.32:

$$\bar{E} = \frac{\beta_o H_o}{\omega\varepsilon} e^{-j\beta z}\hat{x} = \eta H_o e^{-j\beta z} \quad \hat{x} \tag{4.34}$$

where $\eta = \sqrt{\mu_o/\varepsilon}$ is the wave impedance of the medium between the parallel plates. The magnetic field is found in a similar manner from Equation 4.31:

$$\bar{H} = H_o e^{-j\beta z}\hat{y} \tag{4.35}$$

This mode is called the *transverse electromagnetic mode* or TEM mode, because *both the electric and magnetic fields are transverse to the direction of propagation.* Notice that the electromagnetic fields of the TEM mode are *exactly those of a plane wave.* Therefore, the TEM mode in the parallel plate waveguide looks like a plane wave that is truncated at the parallel plates.

An important aspect of the TEM mode is that it is the only mode that can propagate at any frequency down to 0 Hz. Also, the TEM cutoff condition does not depend on the spacing of the waveguide. It can be shown that, in general, the solution of the field amplitudes for a TEM mode are determined by solving the Laplace equation that describes the electrostatic field solution. Electrostatic field solutions only exist if there are at least two conducting surfaces, i.e. it must be possible to apply a voltage potential to generate an electrostatic field. For this reason, a waveguide that supports a TEM mode must have at least two separate conducting surfaces.

The surface current and charge densities on the conductors for the TEM mode can be found from the boundary conditions at a perfectly conducting surface:

$$\hat{n} \cdot \bar{E} = \rho_s/\varepsilon \tag{4.36}$$

$$\hat{n} \times \bar{H} = \bar{J}_s \tag{4.37}$$

where ρ_s is the surface charge density (C/m^2), $\bar{J}_s$ is the surface current (A/m), and $\hat{n}$ is a unit vector normal to the surface and pointing into the dielectric. In the top conductor, the time-varying surface current and surface charge are

$$\bar{J}_{s,\text{top}} = -\hat{x} \times H_o e^{-j\beta z}\hat{y} = -H_o e^{-j\beta z}\hat{z} \tag{4.38}$$

$$\rho_{s,\text{top}} = -\hat{x} \cdot \varepsilon\eta H_o e^{-j\beta z} \quad \hat{x} = -\varepsilon\eta H_o e^{-j\beta z} \tag{4.39}$$

and on the bottom conductor

$$\bar{J}_{s,\text{bott}} = \hat{x} \times H_o e^{-j\beta z}\hat{y} = H_o e^{-j\beta z}\hat{z} \tag{4.40}$$

$$\rho_{s,\text{bott}} = \hat{x} \cdot \varepsilon\eta H_o e^{-j\beta z} \quad \hat{x} = \varepsilon\eta H_o e^{-j\beta z} \tag{4.41}$$

4.3.9 Guide Impedance in a Parallel Plate Waveguide

When using waveguides to transmit electromagnetic waves, transitions from one type or geometry of waveguide to another will often exist. At the interface between the two dissimilar waveguides, reflections will occur. We can determine the reflected and transmitted waves by examining the "guide impedance" of the different waveguide regions.

Consider a parallel plate waveguide composed of two dissimilar sections. The first section of the waveguide is filled with a dielectric having a permittivity ε_a, while the second section is filled with a dielectric ε_b as shown in Figure 4.16. We will first consider the TE case. Due to reflections, an incident and a reflected field will exist in region a, while only a transmitted wave will exist in region b. At the interface between waveguide region a and region b, the total tangential electric and magnetic fields must be continuous. This yields

$$E_{y,i} + E_{y,r} = E_{y,t} \tag{4.42}$$

$$H_{x,i} + H_{x,r} = H_{x,t} \tag{4.43}$$

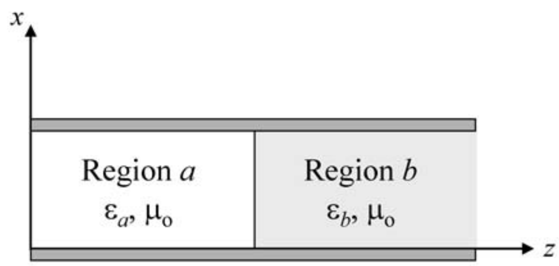


Figure 4.16: A parallel-plate waveguide filled with two different dielectric materials. For a mode propagating in the +z-direction, a reflection will occur at the interface between the two regions.

The power flow in the -z-direction is determined from the cross-product of the y-component of the electric field with the x-component of the magnetic field. For this reason, we define a TE mode guide impedance with respect to power flow in the +z-direction for each region as

$$\eta_g^{\mathrm{TE}} = -\frac{E_{y,i}}{H_{x,i}} = \frac{E_{y,r}}{H_{x,r}} \tag{4.44}$$

Substituting the electric and magnetic field solutions (Equations 4.27 and 4.28) yields a guide impedance of

$$\eta_g^{\mathrm{TE}} = -\frac{E_o \sin\left(\frac{m\pi}{a}x\right) e^{-j\beta z}}{-\frac{\beta E_o}{\omega\mu_o} \sin\left(\frac{m\pi}{a}x\right) e^{-j\beta z}} = \frac{\omega\mu_o}{\beta} = \frac{k\eta}{\beta} \tag{4.45}$$

where $\eta = \sqrt{\mu_o/\varepsilon}$ is the wave impedance of each region and $k = \omega\sqrt{\mu_o\varepsilon}$ is the propagation constant in each region. Using Equation 4.23, the TE guide impedance can be rewritten as

$$\eta_g^{\mathrm{TE}} = \frac{\eta}{\sqrt{1 - f_{c,m}^2/f^2}} = \frac{\eta}{\sqrt{1 - \lambda^2/\lambda_{c,m}^2}} \tag{4.46}$$

where $f_{c,m}$ and $\lambda_{c,m}$ are the cutoff frequency and wavelength of the m^{th} mode, respectively, in region a or region b. For the TM case, continuity of tangential fields gives

$$E_{x,i} + E_{x,r} = E_{x,t} \tag{4.47}$$

$$H_{y,i} + H_{y,r} = H_{y,t} \tag{4.48}$$

We define a TM mode guide impedance for each region as

$$\eta_g^{\mathrm{TM}} = \frac{E_{x,i}}{H_{y,i}} = -\frac{E_{x,r}}{H_{y,r}} \tag{4.49}$$

Substituting the TM field solutions (Equations 4.31 and 4.32) we find

$$\eta_g^{\mathrm{TM}} = \frac{\frac{\beta H_o}{\omega \varepsilon_1} \cos\left(\frac{m\pi}{a}x\right) e^{-j\beta z}}{H_o \cos\left(\frac{m\pi}{a}x\right) e^{-j\beta z}} = \frac{\beta}{\omega\varepsilon} = \frac{\beta\eta}{k} \tag{4.50}$$

Using Equation 4.23, the TM guide impedance can be rewritten as

$$\eta_g^{\mathrm{TM}} = \eta\sqrt{1 - f_{c,m}^2/f^2} = \eta\sqrt{1 - \lambda^2/\lambda_{c,m}^2} \tag{4.51}$$

The analysis of discontinuities in waveguides proceeds exactly as for a plane wave at normal incidence to an interface, provided that we use the guide impedance of each section as opposed to the wave impedance in each dielectric region. The electric field reflection coefficient between waveguide regions 1 and 2 for either a TE or TM mode is

$$\Gamma_{ab} = \frac{\eta_{g,b} - \eta_{g,a}}{\eta_{g,b} + \eta_{g,a}} \tag{4.52}$$

where η_g for regions a and b are calculated using Equation 4.46 or Equation 4.51 depending upon whether we are interested in TE or TM modes, respectively.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 4.3:
Consider a section of air-filled parallel plate waveguide, which is connected to a section filled with a nonmagnetic medium having a permittivity $\varepsilon = 9\varepsilon_o$. Both sections have a spacing between the plates of 10 cm. The frequency of operation is 2 GHz. Find the reflection coefficients of the TE_1 and TM_1 modes at the discontinuity.

Solution:
The cutoff frequency for the $m = 1$ modes in the first (air-filled) region is

$$f_{c,1} = \frac{c}{2a\sqrt{\varepsilon_{\mathrm{r}1}}} = \frac{3 \times 10^8}{2\,(0.10)\,\sqrt{1}} = 1.5 \quad \mathrm{GHz} \quad (\mathrm{region}\, a)$$

and in the second ($\varepsilon = 9\varepsilon_o$) region is

$$f_{c,1} = \frac{c}{2a\sqrt{\varepsilon_{\mathrm{r}2}}} \frac{3 \times 10^8}{2\,(0.10)\,\sqrt{9}} = 0.5 \quad \mathrm{GHz} \quad (\mathrm{region}\, b)$$

Because $f > f_c$ in both cases, the modes can propagate in each region. The reflection coefficient between regions a and b for the TE_1 mode is then found from the guide impedances of the two regions:

$$\eta_{g,a}^{\mathrm{TE}} = \frac{\eta_o}{\sqrt{1 - (1.5/2)^2}} = 1.512\eta_o \quad \Omega$$

$$\eta_{g,b}^{\mathrm{TE}} = \frac{\eta_o/\sqrt{9}}{\sqrt{1 - (0.5/2)^2}} = 0.344\eta_o \quad \Omega$$

$$\Gamma_{ab}^{\mathrm{TE}} = \frac{0.344\eta_o - 1.512\eta_o}{0.344\eta_o + 1.512\eta_o} = -0.629$$

The reflection coefficient for the TM_1 mode is found in a similar manner:

$$\eta_{g,a}^{\mathrm{TM}} = \eta_o\sqrt{1-(1.5/2)^2} = 0.661\eta_o \quad \Omega$$

$$\eta_{g,b}^{\mathrm{TM}} = \left(\frac{\eta_o}{\sqrt{9}}\right)\sqrt{1-(0.5/2)^2} = 0.323\eta_o \quad \Omega$$

$$\Gamma_{ab}^{\mathrm{TM}} = \frac{0.323\eta_o - 0.661\eta_o}{0.323\eta_o + 0.661\eta_o} = -0.343$$

$\ll\ll\ll\ll\ll\ll\ll\gg\gg\gg\gg\gg\gg\gg$

4.4 Symmetric Dielectric Planar Waveguides

The study of planar dielectric waveguides is a good foundation for understanding modes of more complicated waveguide geometries such as optical fibers. This is because, unlike parallel plate waveguides, the modes in a planar dielectric waveguide resemble those in optical fibers. For parallel plate waveguides, determination of the field solutions was fairly straightforward due to the assumption of perfectly conducting walls, and we found field solutions in which the tangential electric field was zero at the walls. For dielectric waveguides, however, the field is guided by reflection at the interface between two dielectric materials, which means that the field solutions extend outside of the guiding region, or "core." This makes the electromagnetic analysis somewhat more complicated for these types of waveguides.

It is important to note that the parallel plate waveguide, which consisted of metallic plates, is useful in the microwave regime, and it is customary to specify the permittivity, ε, and to think in terms of frequency for these types of waveguides. Dielectric waveguides, on the other hand, are useful in the optical regime, and we usually specify the index of refraction, n, and think in terms of wavelength.

The simplest planar waveguide consists of one center planar region of finite width, surrounded by two "infinitely" thick dielectric regions. The outer regions have the same index of refraction, and the inner layer has the highest index of refraction. Because the two outer regions have the same index of refraction, this waveguide is called a "symmetric dielectric planar waveguide." This type of waveguide propagates only TE and TM modes (no TEM mode). The geometry of the symmetric slab waveguide is shown below in Figure 4.17. The center dielectric layer, called the "core" has an index of refraction n_1 and a thickness a. The core is bounded by two semi-infinitely thick layers with index of refraction n_2 collectively called the "cladding." Each of the planar layers are assumed to extend indefinitely in the $\pm$y- and $\pm$z-directions.

4.4.1 Wave Vectors in the Core and Cladding

Light can be "trapped" in the core region of the planar dielectric waveguide if total internal reflection (TIR) occurs and the waveguiding condition is satisfied. Recall that, for TIR to occur, the incident medium must have a higher index of refraction than the second medium. Therefore, for light to be "trapped" in the core of the waveguide, the index of refraction of the core must be greater than the index of refraction of the cladding, i.e. $n_1 > n_2$. For a mode in a dielectric planar waveguide, we can picture the light "bouncing" down the waveguide between the top and bottom interfaces as shown in Figure 4.18. TIR occurs if the "bounce angle," θ, is greater than the critical angle, θ_c, at the core/cladding interface, where

$$\theta_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{4.53}$$

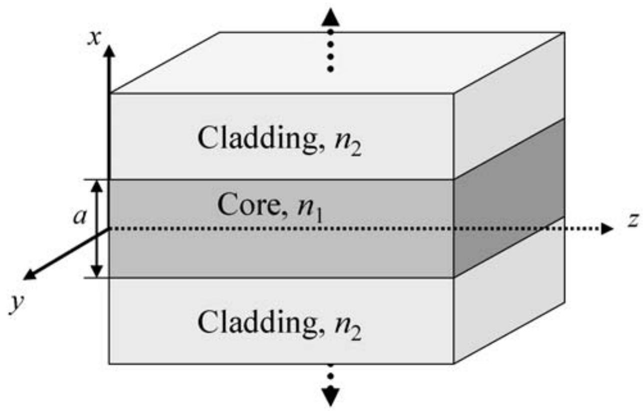


Figure 4.17: Symmetric dielectric planar waveguide. The cladding layers are assumed to be "infinitely" thick. The cladding and core layers extend out to infinity in the ±y- and ±z-directions. Note that, in contrast to the parallel plate case, the z-axis is defined a running through the *center* of the core. Because of symmetry, this simplifies the field solutions. This does not affect the waveguiding condition derived in the previous chapter.

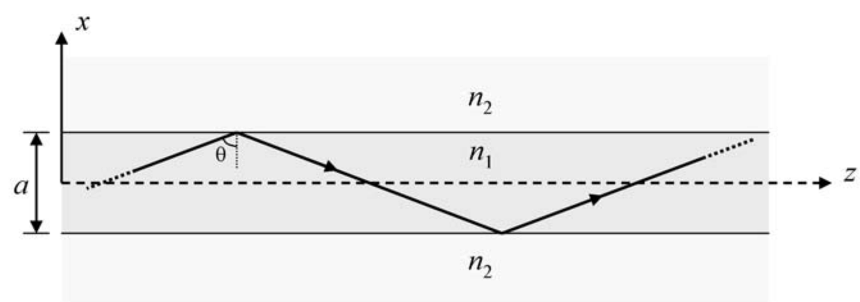


Figure 4.18: A symmetric dielectric slab waveguide guides light by TIR. Note that, in contrast to the parallel plate case, the z-axis is defined as running through the *center* of the core. Because of symmetry, this simplifies the field solutions. This does not affect the waveguiding condition derived in the previous chapter.

Without loss of generality, we confine the wave vectors of the "bouncing" wave to the x-z plane, as shown in Figure 4.18. The wave vector in the core has a direction determined by the direction of the bouncing ray, and has a magnitude $k_1 = n_1 k_o$. Likewise, the magnitude of the wave vector in the cladding is $k_2 = n_2 k_o$. The relationships between the components of the wave vectors in the core and cladding are shown below in Figure 4.19.

The next question we ask is, without going into the waveguiding condition, is it possible to place limits on the values of β that can correspond to guided modes? In the core region, the wave vector components are related by

$$(n_1 k_o)^2 = \kappa_1^2 + \beta^2 \tag{4.54}$$

where $\beta = n_1 k_o \sin(\theta)$ and $\kappa_1 = n_1 k_o \cos(\theta)$ (see Figure 4.19). The maximum possible value for β occurs when the wave is propagating directly down the waveguide ($\theta = 90°$). In this case, $\kappa_1 = 0$ and $\beta = n_1 k_o$. The minimum possible value of β that still corresponds to a guided wave occurs when the wave vector is at

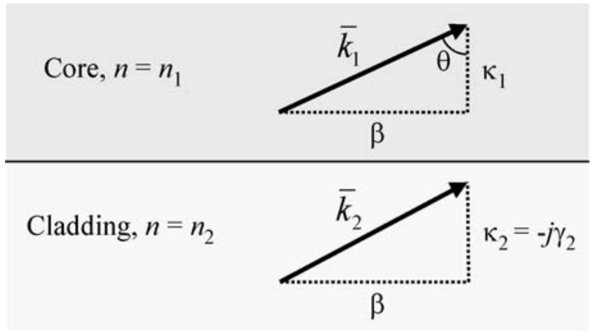


Figure 4.19: Relationships between the wave vector components in the core and cladding regions of the waveguide. Note that the upper cladding is not shown in this figure. Recall that β must be equal in the core and cladding regions.

the critical angle, since θ must be greater than θ_c for TIR to occur. If $\theta = \theta_c$, then we have

$$\beta = n_1 k_o \sin(\theta_c) = n_1 k_o \left(\frac{n_2}{n_1}\right) = n_2 k_o \tag{4.55}$$

Therefore, the propagation constant of a guided mode must satisfy

$$n_2 k_o < \beta < n_1 k_o \tag{4.56}$$

Defining the *effective index* of the mode as $n_{\text{eff}} = \beta/k_o$, this condition can be rewritten as

$$n_2 < n_{\text{eff}} < n_1 \tag{4.57}$$

What conclusions can we now draw about the wave vector components in the cladding? In the cladding region, the wave vector components are related by

$$(n_2 k_o)^2 = \kappa_2^2 + \beta^2 \tag{4.58}$$

Recall that the tangential wave vector component (β in this case) must be continuous across a dielectric interface (see Chapter 3). Therefore β must be equal in the core and the cladding regions. Solving for the x-component of the wave vector in the cladding, κ_2, we find

$$\kappa_2 = \sqrt{(n_2 k_o)^2 - \beta^2} \tag{4.59}$$

However, for waveguiding to occur it must be true that $\beta > n_2 k_o$. This means that in the cladding, κ_2 is an *imaginary number*, so we define

$$\kappa_2 \equiv -j\gamma_2 \tag{4.60}$$

such that

$$\gamma_2 = \sqrt{\beta^2 - (n_2 k_o)^2} \tag{4.61}$$

4.4.2 Propagation Constants of the Guided Modes

In the previous section, we established that the effective index of modes guided by the symmetric planar waveguide must lie between the index of the core and the cladding. Does this mean that any value of n_{eff} (i.e. any bounce angle) that meets this condition corresponds to a guided mode? The answer is no, since the waveguiding condition must also be satisfied. To investigate this, we return to the waveguiding condition (Equation 4.12):

$$r_{10} r_{12} e^{-j2a\kappa_1} = 1 \tag{4.62}$$

Because we are considering only a symmetric structure, the reflection coefficients from the dielectric interfaces, r_{10} and r_{12}, will be equal, and we define $r_{10} = r_{12} \equiv r$. In Chapter 3 we found that the reflection coefficient at a dielectric interface is

$$r = \frac{\kappa_1 - \delta_{12}\kappa_2}{\kappa_1 + \delta_{12}\kappa_2} = \frac{\kappa_1 + j\delta_{12}\gamma_2}{\kappa_1 - j\delta_{12}\gamma_2} \equiv e^{j2\psi} \tag{4.63}$$

where $\delta_{12} = 1$ for TE polarization or $\delta_{12} = \varepsilon_1/\varepsilon_2$ for TM polarization, and we have defined

$$\psi = \tan^{-1}\left(\delta_{12}\frac{\gamma_2}{\kappa_1}\right) \tag{4.64}$$

Notice that the magnitude of r is unity, which confirms that TIR is occurring. Also note that a phase shift of 2ψ will occur with each reflection from the core/cladding interface. Substituting the reflection coefficient back into Equation 4.62, the waveguiding condition becomes

$$e^{j(4\psi - 2\kappa_1 a)} = 1 \tag{4.65}$$

for the symmetric dielectric slab waveguide. This equation is satisfied only when

$$4\psi - 2\kappa_1 a = -2\pi\nu \tag{4.66}$$

where $\nu = 0, 1, 2, ...$ Because each bounce angle corresponds to a unique value of κ_1, this equation shows that *only certain bounce angles correspond to guided modes*! Substituting Equation 4.64 into Equation 4.66 we obtain

$$\tan^{-1}\left(\delta_{12}\frac{\gamma_2}{\kappa_1}\right) = \frac{\kappa_1 a}{2} - \frac{\nu\pi}{2} \tag{4.67}$$

This is the transcendental equation that determines the allowed modes in a symmetric dielectric slab waveguide. It is customary to define normalized parameters to simplify the solution of the equation. The *normalized frequency*, or *V-number*, is defined as

$$V = \frac{\pi a}{\lambda_o}\sqrt{n_1^2 - n_2^2} = \frac{k_o a}{2}\sqrt{n_1^2 - n_2^2} \tag{4.68}$$

and the *normalized propagation constant*, b, is defined as

$$b = \frac{(\beta/k_o)^2 - n_2^2}{n_1^2 - n_2^2} = \frac{n_{\text{eff}}^2 - n_2^2}{n_1^2 - n_2^2} \tag{4.69}$$

The requirement that $n_2 < n_{\text{eff}} < n_1$ means that the normalized propagation constant, b, must lie between 0 and 1. Using these normalized parameters, it can be shown that

$$\frac{\gamma_2}{\kappa_1} = \frac{\sqrt{\beta^2 - k_2^2}}{\sqrt{k_1^2 - \beta^2}} = \sqrt{\frac{(\beta/k_o)^2 - n_2^2}{n_1^2 - (\beta/k_o)^2}} = \sqrt{\frac{b}{1-b}} \tag{4.70}$$

and

$$\frac{\kappa_1 a}{2} = \frac{a}{2}\sqrt{k_1^2 - \beta^2} = \frac{k_o a}{2}\sqrt{n_1^2 - n_{\text{eff}}^2} = V\sqrt{1-b} \tag{4.71}$$

Therefore, the transcendental equation (Equation 4.67) can be rewritten in terms of the normalized parameters as

$$\tan^{-1}\left(\delta_{12}\sqrt{\frac{b}{1-b}}\right) = V\sqrt{1-b} - \frac{\nu\pi}{2} \tag{4.72}$$

where ν is referred to as the "order" of the mode, and is similar to the mode number m for parallel plate waveguides. Unlike the parallel plate waveguide, however, $\nu \geq 0$ for both TE and TM modes (since there is no TEM mode).

By substituting specific values for b, which must lie between 0 and 1, we can easily solve for the corresponding values of V for a given value of ν. This generates the b-V graph shown in Figure 4.20. Using this graph, it is easy to determine the number of TE modes and the propagation constant of each mode given the V-number of the waveguide. For example, if $V = 4$, the waveguide will support three different TE modes: TE_2, TE_1, and TE_0, where the subscript refers to the order of the mode, ν. The corresponding values of b can be read from the graph: $b \approx 0.20$, 0.62, and 0.90. More accurate solutions may be obtained from tabulated values of b vs. V. If the indices of refraction of the core and cladding are known, we can use these values of b to approximate the propagation constant of each mode, β. Note that the curves shown in Figure 4.20 also approximately apply for TM modes if $n_1 \approx n_2$. This situation occurs quite frequently in waveguides, and is called the "weakly guiding condition."

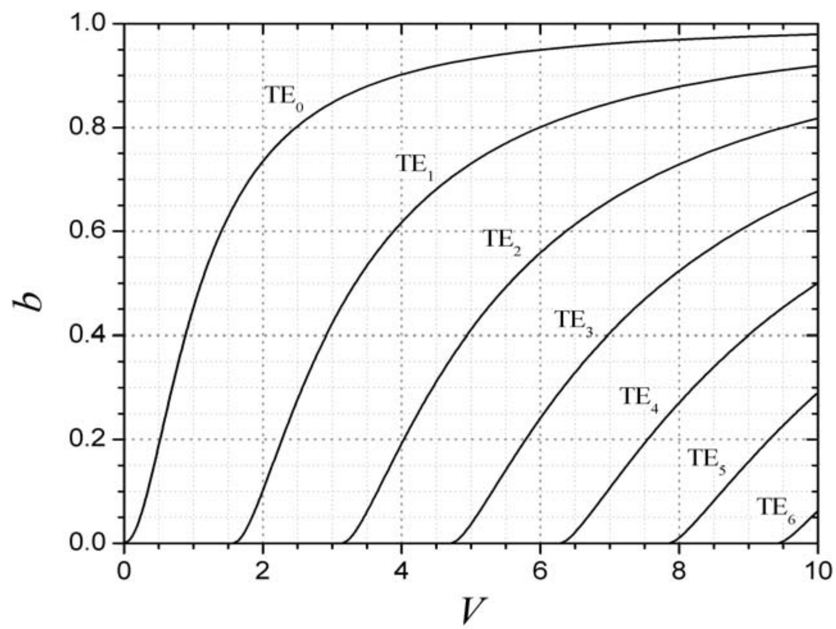


Figure 4.20: Normalized propagation constant b versus normalized frequency V for the TE modes in a symmetric dielectric planar waveguide. The most left curve is the $\nu = 0$ mode, the next curve to the right the $\nu = 1$ mode, and so on. The curves also apply for TM modes if $n_1 \approx n_2$. The curves are calculated by varying b from 0 to 1 to calculate $V = [\tan^{-1}(\delta_{12}\sqrt{b/(1-b)}) + \nu\pi/2]/\sqrt{1-b}$.

Notice that increasing V has two effects. First, the normalized propagation constant, b, of each mode approaches 1. From Equation 4.69, we see that this means that the effective index of each mode, n_{eff}, approaches the index of the core, n_1. Given what you know about TIR, why is this the case? The second effect of increasing V is that the number of modes the waveguide can support increases.

4.4.3 Cutoff Conditions for Guided Modes

For each mode, there is a certain value of V below which the mode is no longer guided (when $b = 0$). Similar to parallel plate waveguides, the terminology for this is that the mode is "cut off" at that particular value of V. The cutoff V-number for each mode, V_c, can be found by substituting $b = 0$ into Equation 4.72, which yields

$$V_{c,\nu} = \frac{\nu\pi}{2} \tag{4.73}$$

This cutoff condition applies for *TE and TM modes*. For example, if $\pi/2 < V \leq \pi$, two TE and two TM modes can propagate in the waveguide. If $\pi < V \leq 3\pi/2$, three TE and three TM modes can propagate, and so on. Returning to the bouncing wave model, the cutoff condition occurs when the wave is at the critical angle. Beyond the critical angle, the wave is not completely reflected by the core/cladding boundary and the mode cannot propagate down the waveguide without loss.

Notice that the cutoff for the $\nu = 0$ mode is $V_{c,0} = 0$. Therefore, the $\nu = 0$ mode *can always propagate in a symmetric slab waveguide*. This means that very low frequency (long wavelength) signals, approaching 0 Hz, can propagate in a *symmetric* dielectric waveguide. In practice, however, to get a perfectly symmetric waveguide is impossible and the cladding regions will have slightly different indices of refraction. For this reason, the cutoff frequency is never at 0 Hz, and in fact, is considerably higher and often in the mid-infrared region.

Because all planar dielectric waveguides support the $\nu = 0$ TE and TM modes, they are called the *fundamental TE* and *fundamental TM* modes. As long as $V \leq \pi/2$, only the fundamental TE and TM modes are allowed to propagate in the waveguide. If this condition is met, the waveguide is referred to as a *single-mode waveguide*. Single-mode waveguides are extremely important in optical communications because they allow signals to be transmitted over long distances with minimal distortion, as we will see later. For a symmetric dielectric planar waveguide the *normalized single-mode cutoff frequency*, V_c, is

$$V_c = \pi/2 \approx 1.571 \tag{4.74}$$

Substituting this value into Equation 4.68 and rearranging, we find the *single-mode cutoff wavelength* for a symmetric dielectric slab waveguide

$$\lambda_c = 2a\sqrt{n_1^2 - n_2^2} \tag{4.75}$$

(Remember that λ_c is the wavelength in free space, not inside the waveguide!) At wavelengths shorter than λ_c, multiple TE and TM modes are allowed to propagate in the waveguide, i.e. the waveguide is a *multimode guide*.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 4.4:

Consider a dielectric planar waveguide with a core width, a, of 1 μm and core and cladding indices of $n_1 = 1.5$ and $n_2 = 1.4$, respectively. Find the single-mode cutoff wavelength, and the number of modes this waveguide will support at a free space wavelength of 0.6 μm.

Solution:

The cutoff wavelength of this waveguide is

$$\lambda_c = 2a\sqrt{n_1^2 - n_2^2} = 2\left(1 \times 10^{-6}\right)\sqrt{1.5^2 - 1.4^2} = 1.08 \ \ \mu\text{m}$$

That means that this particular waveguide will be single-mode for $\lambda_o > 1.08$ μm. If the waveguide is used at a wavelength of 0.6 μm, the waveguide will be multimode with a V parameter of

$$V = \frac{\pi a}{\lambda_o}\sqrt{n_1^2 - n_2^2} = \frac{\pi\left(1 \times 10^{-6}\right)}{0.6 \times 10^{-6}}\sqrt{1.5^2 - 1.4^2} = 2.82$$

From Figure 4.20 we see that the waveguide will support 2 TE and 2 TM modes at that wavelength.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

4.4.4 Electromagnetic Field Analysis

In the previous section, we developed the transcendental equation that determines the modal propagation constants of the dielectric planar waveguide without considering the electromagnetic fields associated with the modes. In this section, we directly solve for the fields associated with the various modes by first solving the wave equation in the core and cladding regions, and then applying the appropriate boundary conditions at the interfaces. We will first derive the fields of the TE modes, followed by the TM modes.

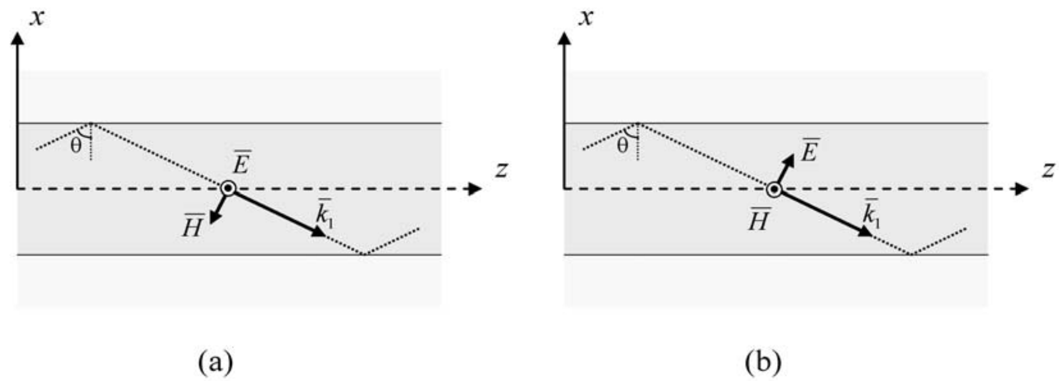


Figure 4.21: Orientation of the electric and magnetic fields for (a) the TE modes, and (b) the TM modes of a dielectric planar waveguide.

Figures 4.21(a) and (b) indicate the direction of the field vectors for TE and TM modes, respectively. Recall that the electric and magnetic field vectors are perpendicular to each other, and both are perpendicular to the direction of propagation (the direction of the wave vector). For TE modes, the electric field is parallel to the core/cladding interface, i.e. in the y-direction. From the figure, we can see that the magnetic field vector of the TE modes will have a components in the x- and z-directions. For the TM modes, the magnetic field vector will be aligned with the y-axis, and the electric field will have components in the x- and z-directions. Therefore, we will expect solutions of the following form for the TE and TM modes:

$$\text{TE modes}: \quad \begin{cases} \bar{E} = E_y \hat{y} \\ \bar{H} = H_x \hat{x} + H_z \hat{z} \end{cases}$$

$$\text{TM modes}: \quad \begin{cases} \bar{E} = E_x \hat{x} + E_z \hat{z} \\ \bar{H} = H_y \hat{y} \end{cases}$$

4.4.5 TE Modes

Because the planar dielectric waveguide is invariant and of infinite extent in the y-direction, we can assume that the fields do not vary in the y-direction. We then assume an electric field solution of the form

$$\bar{E} = E_y(x) e^{-j\beta z} \hat{y} \tag{4.76}$$

In other words, we are assuming propagation in the z-direction with the propagator $e^{-j\beta z}$. When this field solution is substituted into the wave equation ($\nabla^2 \bar{E} + k^2 \bar{E} = 0$), it can be rearranged into the form (see Appendix D):

$$\nabla_t^2 \bar{E} + \left(k^2 - \beta^2\right) \bar{E} = 0 \tag{4.77}$$

where the *transverse Laplacian* operator, $\nabla_t^2 \bar{E}$, is

$$\nabla_t^2 \bar{E} = \frac{\partial^2 \bar{E}}{\partial x^2} + \frac{\partial^2 \bar{E}}{\partial y^2} \tag{4.78}$$

Since we have assumed no field variation in the y-direction, we can conclude that $\partial^2 \bar{E}/\partial y^2 = 0$. From Equation 4.54 we know that $k_1^2 - \beta^2 = \kappa_1^2$ in the core region. Therefore, in the core region the wave equation reduces to

$$\frac{\partial^2}{\partial x^2} E_y(x) + \kappa_1^2 E_y(x) = 0 \qquad (|x| \leq a/2) \tag{4.79}$$

The solutions to this equation are sines and cosines, which means that the field will have sinusoidal behavior in the transverse direction in the core.

In the cladding region the wave vector components are related by $k_2^2 - \beta^2 = (-j\gamma_2)^2$, as given in Equation 4.59. The wave equation in the cladding is therefore

$$\frac{\partial^2}{\partial x^2} E_y(x) - \gamma_2^2 E_y(x) = 0 \qquad (|x| \geq a/2) \tag{4.80}$$

The solution to this equation is an exponential that decays away from the core (an exponential that increases away from the core is also a solution to Equation 4.80; however, this is a non-physical solution since the fields must decay to zero far from the core). Due to the symmetry of the waveguide structure, we know that the modes will either be perfectly symmetric or perfectly antisymmetric about the z-axis. This allows us to write the field solutions in the top cladding, core, and lower cladding:

$$\bar{E} = E_o e^{-\gamma_2(x-a/2)} e^{-j\beta z} \hat{y} \qquad (x \geq a/2) \tag{4.81}$$

$$\bar{E} = \left\{ \begin{array}{c} \cos(\kappa_1 x) \\ \sin(\kappa_1 x) \end{array} \right\} e^{-j\beta z} \hat{y} \qquad (-a/2 \leq x \leq a/2) \tag{4.82}$$

$$\bar{E} = \left\{ \begin{array}{c} + \\ - \end{array} \right\} E_o e^{+\gamma_2(x+a/2)} e^{-j\beta z} \hat{y} \qquad (x \leq -a/2) \tag{4.83}$$

(Refer to Appendix D for an alternate derivation of the general form of the fields.) In the core, the cosine solution corresponds to the even "symmetric" modes (ν = 0, 2, 4, ...) while the sine solution corresponds to the odd "antisymmetric" modes (ν = 1, 3, 5, ...). Notice that we have arbitrarily normalized the field in the core region to have a maximum amplitude of unity. The values of E_o that correspond to guided modes are determined by applying the boundary conditions at the core/cladding interface. In Chapter 3 we saw that the tangential electric and magnetic fields must be continuous across a dielectric interface. Applying the continuity of the *tangential electric fields* at $x = a/2$ (i.e. setting Equation 4.81 equal to Equation 4.82 at $x = a/2$) yields

$$E_o = \left\{ \begin{array}{c} \cos(\kappa_1 a/2) \\ \sin(\kappa_1 a/2) \end{array} \right. \tag{4.84}$$

Note that we can also apply the boundary conditions at $x = -a/2$, but this redundant as it yields the same equations.

Referring to the field solutions (Equations 4.81-4.83), we observe that the solutions in the core are standing waves with varying periods depending upon the value of κ_1. In the cladding regions, the solutions are decaying, or *evanescent*, fields with a decay rate depending upon the value of γ_2. The evanescent "tail" decays toward zero in either direction away from the core region. The positive and negative sign in the lower cladding solution correspond to the symmetric and antisymmetric solutions, respectively.

Using Faraday's Law, it is straightforward to find the magnetic fields associated with the TE modes of the planar slab waveguide:

$$\bar{H} = \frac{j}{\omega\mu_o} \nabla \times \bar{E} = \frac{j}{\omega\mu_o} \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & E_y & 0 \end{vmatrix} = \frac{j}{\omega\mu_o} \left(-\frac{\partial E_y}{\partial z} \hat{x} + \frac{\partial E_y}{\partial x} \hat{z} \right) \tag{4.85}$$

Evaluating this equation gives the magnetic fields for the upper cladding region ($x \geq a/2$):

$$\bar{H} = \frac{E_o}{\omega\mu_o}\left(-\beta\hat{x} - j\gamma_2\hat{z}\right)e^{-\gamma_2(x-a/2)}e^{-j\beta z} \tag{4.86}$$

the core region ($-a/2 \leq x \leq a/2$):

$$\bar{H} = \frac{1}{\omega\mu_o}\left(-\beta\left\{\begin{array}{c}\cos(\kappa_1 x)\\ \sin(\kappa_1 x)\end{array}\right\}\hat{x} - j\kappa_1\left\{\begin{array}{c}\sin(\kappa_1 x)\\ -\cos(\kappa_1 x)\end{array}\right\}\hat{z}\right)e^{-j\beta z} \tag{4.87}$$

and the lower cladding region ($x \leq a/2$):

$$\bar{H} = \left\{\begin{array}{c}+\\ -\end{array}\right\}\frac{E_o}{\omega\mu_o}\left(-\beta\hat{x} + j\gamma_2\hat{z}\right)e^{+\gamma_2(x+a/2)}e^{-j\beta z} \tag{4.88}$$

where the upper solutions again refer to the symmetric (even) modes, and the lower solutions correspond to the antisymmetric (odd) modes.

By solving the waveguiding condition (Equation 4.72) and substituting the propagation constant, β, back into the field equations we obtain the electric field profiles of the modes. The electric field profiles of the four lowest-order TE modes are shown in Figure 4.22. Notice that the TE_0 mode has a single maximum at the center of the core and no nulls (zero crossings). The TE_1 mode has two maxima and one null. In general, the mode number indicates the number of nulls for a particular mode.

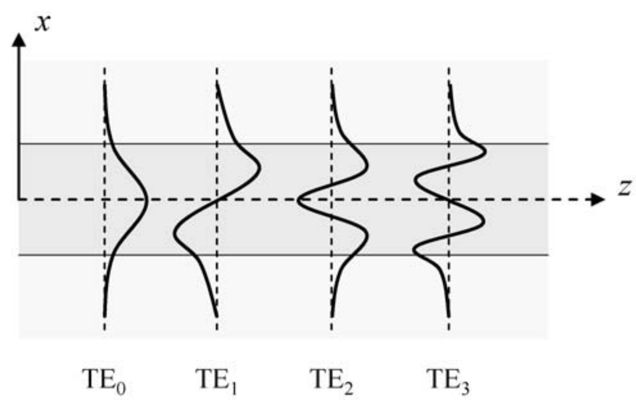


Figure 4.22: The electric field profile of the four lowest-order TE modes in a symmetric dielectric planar waveguide.

Let us now examine a mode when it approaches cutoff as the wavelength is increased (V, which is proportional to frequency, is decreased). Figure 4.23 shows the evolution of the TE_1 mode as the wavelength is increased (V is decreased). From the b-V curve (Figure 4.20), we see that at shorter wavelengths (large V-numbers), the effective index of the mode, n_{eff}, approaches the index of the core, n_1. From the field solutions, we see that in this case the electric field is tightly confined within the core. As the wavelength is increased (V is decreased), n_{eff} approaches the index of the cladding, and we move toward the critical angle for TIR. This causes the effective index to approach n_2. The attenuation constant of the fields in the cladding decreases and the mode is less strongly confined in the core.

It should be noted that, even though there is electromagnetic power propagating in the cladding, all of the power is still propagating in the z-direction, i.e. there is no light "leaking" out of the waveguide as the mode propagates. However, once the wavelength is increased sufficiently such that cutoff occurs, then the solutions in the cladding become propagating waves in both the x- and z-directions (β and κ_2 are real). In this case, we no longer have a waveguide that only propagates power in the z-direction.

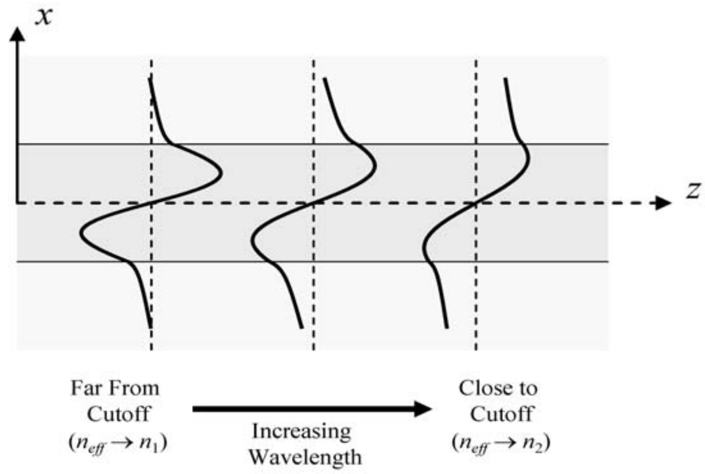


Figure 4.23: As a mode approaches cutoff, more of the field propagates in the cladding, which causes the effective index of the mode to decrease.

Another way to picture the evolution of the mode as the wavelength is changed is with the β-ω graph shown in Figure 4.24. In this graph, the propagation constant of each mode, β, is plotted against the *frequency* of the light propagating in the waveguide (recall that $\omega = 2\pi c/\lambda_o$). At cutoff, the propagation constant of each mode is equal to the propagation constant in the cladding, $\omega n_2/c$. As the frequency is increased, the propagation constant approaches the propagation constant of the core, $\omega n_1/c$. The most important point to notice from the β-ω graph is that the curves for each mode are *not straight lines*! This is true for all types of dielectric waveguides, including optical fibers. While the meaning of this is unclear now, it will become very important when we discuss *waveguide dispersion*, which causes pulses of light to distort as they travel down the waveguide.

$\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 4.5:

For the waveguide described in Example 4.4, find the electric field of the TE_1 mode at a free space wavelength of 0.6 μm.

Solution:

In Example 4.4, we found that $V = 2.82$ for this waveguide at $\lambda_o = 0.6$ μm. From Figure 4.20 (or from a table of b vs. V values), we see that this corresponds to a normalized propagation constant of $b \approx 0.38$ for the TE_1 mode. The free space propagation constant is

$$k_o = \frac{2\pi}{0.6 \times 10^{-6}} = 1.05 \times 10^7 \quad \text{rad/m}$$

Solving Equation 4.69 for β we find

$$\beta = k_o\sqrt{n_2^2 + b(n_1^2 - n_2^2)} = 1.05 \times 10^7 \sqrt{1.4^2 + 0.38\,(1.5^2 - 1.4^2)} = 1.51 \times 10^7 \quad \text{rad/m}$$

The transverse component of the wave vector in the core is

$$\kappa_1 = \sqrt{(n_1 k_o)^2 - \beta^2} = \sqrt{(1.5 \times 1.05 \times 10^7)^2 - (1.51 \times 10^7)^2} = 4.44 \times 10^6 \quad \text{rad/m}$$

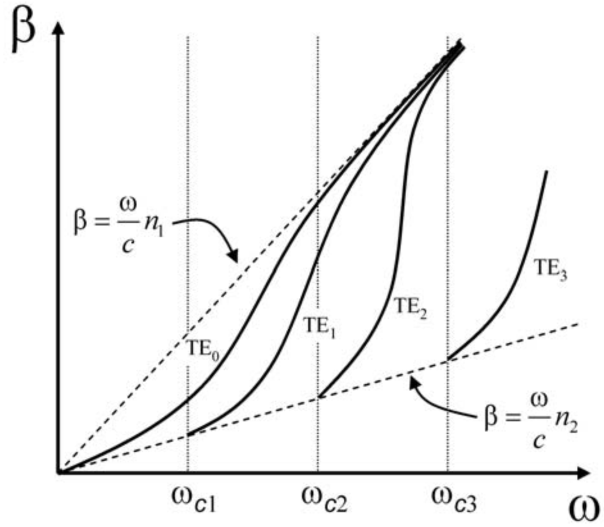


Figure 4.24: β-ω graph for the three lowest-order TE modes in a symmetric dielectric planar waveguide.

The transverse attenuation constant in the cladding is

$$\gamma_2 = \sqrt{\beta^2 - (n_2 k_o)^2} = \sqrt{(1.51 \times 10^7)^2 - (1.4 \times 1.05 \times 10^7)^2} = 3.48 \times 10^6 \quad \text{Np/m}$$

The constant E_o is then (note that this is an odd mode)

$$E_o = \sin\left(\frac{\kappa_1 a}{2}\right) = \sin\left\{\frac{(4.44 \times 10^6)(1 \times 10^{-6})}{2}\right\} = 0.796$$

Inserting these values into the field solutions for an odd mode we find:

$$\bar{E} = 0.796 e^{-3.48\times10^6(x-0.5\times10^{-6})} e^{-j1.51\times10^7 z}\hat{y} \qquad (x \geq 0.5\mu m)$$

$$\bar{E} = \sin(4.44 \times 10^6 x)\, e^{-j1.51\times10^7 z}\hat{y} \qquad (-0.5\mu m \leq x \leq 0.5\mu m)$$

$$\bar{E} = -0.796 e^{+3.48\times10^6(x+0.5\times10^{-6})} e^{-j1.51\times10^7 z}\hat{y} \qquad (x \leq -0.5\mu m)$$

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

4.4.6 TM Modes

The derivation of the TM modes proceeds in an analogous manner, so only the results will be presented here. For the TM modes, we begin with the magnetic fields in the core and cladding:

$$\bar{H} = H_o e^{-\gamma_2(x-a/2)} e^{-j\beta z} \; \hat{y} \qquad (x \geq a/2) \tag{4.89}$$

$$\bar{H} = \left\{ \begin{array}{c} \cos(\kappa_1 x) \\ \sin(\kappa_1 x) \end{array} \right\} e^{-j\beta z} \; \hat{y} \qquad (-a/2 \leq x \leq a/2) \tag{4.90}$$

$$\bar{H} = \left\{ \begin{array}{c} + \\ - \end{array} \right\} H_o e^{+\gamma_2 (x+a/2)} e^{-j\beta z} \hat{y} \qquad (x \leq -a/2) \tag{4.91}$$

Using Ampere's Law, we find the electric fields associated with the TM modes of the planar slab waveguide:

$$\bar{E} = \frac{-j}{\omega\varepsilon} \nabla \times \bar{H} = \frac{-j}{\omega\varepsilon} \left| \begin{array}{ccc} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ 0 & H_y & 0 \end{array} \right| = \frac{j}{\omega\varepsilon} \left(\frac{\partial H_y}{\partial z} \hat{x} - \frac{\partial H_y}{\partial x} \hat{z} \right) \tag{4.92}$$

where $\varepsilon = n_1^2 \varepsilon_o$ in the core and $\varepsilon = n_2^2 \varepsilon_o$ in the cladding. Evaluating this equation gives the electric fields for the upper cladding region ($x \geq a/2$):

$$\bar{E} = \frac{H_o}{\omega n_2^2 \varepsilon_o} \left(\beta \hat{x} + j\gamma_2 \hat{z} \right) e^{-\gamma_2 (x-a/2)} e^{-j\beta z} \tag{4.93}$$

the core region ($-a/2 \leq x \leq a/2$):

$$\bar{E} = \frac{1}{\omega n_1^2 \varepsilon_o} \left(\beta \left\{ \begin{array}{c} \cos(\kappa_1 x) \\ \sin(\kappa_1 x) \end{array} \right\} \hat{x} + j\kappa_1 \left\{ \begin{array}{c} \sin(\kappa_1 x) \\ -\cos(\kappa_1 x) \end{array} \right\} \hat{z} \right) e^{-j\beta z} \tag{4.94}$$

and the lower cladding region ($x \leq a/2$):

$$\bar{E} = \left\{ \begin{array}{c} + \\ - \end{array} \right\} \frac{H_o}{\omega n_2^2 \varepsilon_o} \left(+\beta \hat{x} - j\gamma_2 \hat{z} \right) e^{+\gamma_2 (x+a/2)} e^{-j\beta z} \tag{4.95}$$

where the upper solutions again refer to the symmetric modes, and the lower solutions correspond to the antisymmetric modes. Applying the continuity of the *tangential magnetic fields* at $x = a/2$ yields

$$H_o = \left\{ \begin{array}{c} \cos(\kappa_1 a/2) \\ \sin(\kappa_1 a/2) \end{array} \right. \tag{4.96}$$

4.5 Rectangular Waveguides

In the previous sections, we studied planar waveguides, in which the field is confined only in the x-direction. In this section, we will study rectangular waveguides. In these types of waveguides, the field is confined in the x- and y-directions with conducting plates. The geometry of a rectangular waveguide is shown in Figure 4.25. The waveguide has a width a in the x-direction and a height b in the y-direction By convention $a \geq b$. We also assume that the waveguide is of infinite extent in the z-direction. The waveguide is filled with a nonmagnetic material with permittivity ε, and we assume that the walls are perfectly conducting.

4.5.1 Propagation Constants of the Guided Modes

We can derive the propagation constants of the guided modes without undertaking a detailed electromagnetic analysis. Because we have placed boundaries in the x- and y-directions, we can no longer assume propagation only in the x-z plane. Instead, a wave propagates down the rectangular waveguide by "bouncing" off all of the walls, and therefore the wave vector inside the guide may have components in the x-, y-, and/or z-directions. Therefore, we decompose the wave vector in the guiding region into x-, y-, and z-components:

$$\bar{k}_1 = k_x \hat{x} + k_y \hat{y} + \beta \hat{z} \tag{4.97}$$

where the magnitudes of the components are related by

$$k_1^2 = \omega^2 \mu_o \varepsilon = k_x^2 + k_y^2 + \beta^2 \tag{4.98}$$

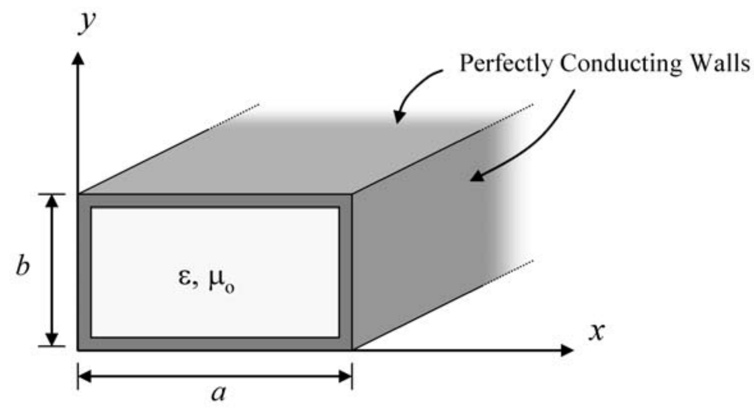


Figure 4.25: Geometry of a rectangular waveguide. The walls are assumed to be perfectly conducting, and by convention $a \geq b$.

We can think of the rectangular waveguide as two parallel-plate waveguides: one in the x-direction and one in the y-direction. Therefore, the waveguiding condition for the parallel-plate waveguide, Equation 4.14, still applies in the x- and y-directions. In other words, the value of k_x will be restricted such that

$$ak_x = m\pi \tag{4.99}$$

where $m = 0, 1, 2, ...$ is the mode number in the x-direction. As in the case of the parallel plate waveguide, the number m indicates the number of maxima/minima in the x-direction. In a similar manner, values of k_y will be restricted by the walls so that

$$bk_y = n\pi \tag{4.100}$$

where $n = 0, 1, 2, ...$ is the mode number in the y-direction and indicates the number of maxima/minima in that direction (note that by convention n is used as the mode number, and does *not* correspond to the index of refraction). From Equation 4.98 we then find that the propagation constant for a particular combination of m and n is

$$\beta_{mn} = \sqrt{\omega^2 \mu_o \varepsilon - (m\pi/a)^2 - (n\pi/b)^2} \tag{4.101}$$

The corresponding modes of the rectangular waveguide are referred to as TE_{mn} or TM_{mn} modes, where $E_z = 0$ (the electric fields lie in the xy-plane) for TE modes and $H_z = 0$ (the magnetic fields lie in the xy-plane) for TM modes. Since by convention $a \geq b$, the first subscript refers to the mode number along the *larger dimension of the rectangular waveguide.*

As in the case of the parallel plate waveguide, the boundary conditions restrict certain modes from propagating. The following modes *cannot exist in a rectangular waveguide*:

$$\text{TE}_{00}, \quad \text{and} \quad \text{TM}_{00}, \text{TM}_{m0}, \text{TM}_{0n}$$

Given the boundary conditions at a perfectly conducting interface, why are these modes restricted from propagating?

4.5.2 Cutoff Conditions

Recall that a mode enters cutoff when the propagation constant becomes imaginary. To find the cutoff frequency for a given mode in a rectangular waveguide, we set β_{mn} equal to zero and solve for the frequency. This yields

$$f_{c,mn} = \frac{1}{2\sqrt{\mu_o \varepsilon}} \sqrt{(m/a)^2 + (n/b)^2} = \frac{c}{2\sqrt{\varepsilon_r}} \sqrt{(m/a)^2 + (n/b)^2} \tag{4.102}$$

where ε_r is the relative permittivity of the filling material. Or, in terms of wavelength

$$\lambda_{c,mn} = \frac{2\sqrt{\varepsilon_r}}{\sqrt{(m/a)^2 + (n/b)^2}} \tag{4.103}$$

Note that $\lambda_{c,mn}$ is the wavelength in *free space*, not the wavelength in the waveguide.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

EXAMPLE 4.6:

Consider a rectangular metallic waveguide with dimensions $a = 10$ cm and $b = 5$ cm, filled with a dielectric with a permittivity $\varepsilon = 4\varepsilon_o$. Find the cutoff frequencies of the 4 lowest-order TE modes. How many TE and TM modes can propagate at 1 GHz?

Solution:

We know that the TE_{00} mode cannot exist in a rectangular waveguide. From Equation 4.102 we see that the four modes with the lowest cutoff frequencies are the TE_{10}, TE_{01}, TE_{20}, and TE_{11} modes:

$$f_{c,10} = \frac{3 \times 10^8}{2\sqrt{4}}\sqrt{(1/0.1)^2} = 0.75\,\text{GHz}$$

$$f_{c,01} = \frac{3 \times 10^8}{2\sqrt{4}}\sqrt{(1/0.05)^2} = 1.50\,\text{GHz}$$

$$f_{c,20} = \frac{3 \times 10^8}{2\sqrt{4}}\sqrt{(2/0.1)^2} = 1.50\,\text{GHz}$$

$$f_{c,11} = \frac{3 \times 10^8}{2\sqrt{4}}\sqrt{(1/0.1)^2 + (1/0.05)^2} = 1.68\,\text{GHz}$$

What is the lowest-order TM mode that can propagate in this rectangular waveguide? Since the TM_{00}, TM_{m0}, and TM_{0n} modes cannot exist, the lowest-order TM mode is the TM_{11} mode, which has a cutoff frequency of 1.68 GHz. Therefore at 1 GHz only the TE_{10} mode can propagate and the waveguide is single-mode, since $f > f_c$ only for that mode. Notice that since $a = 2b$ the TE_{01} and TE_{20} modes have the same cutoff frequency. The significance of this will become apparent later in this chapter. The propagation constants of the 5 lowest-order modes in this waveguide are shown in Figure 4.26.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

4.5.3 Electromagnetic Modes

In general, in a rectangular waveguide a variety of TE and TM modes may exist, and the field expressions are fairly complex. Therefore, we will not present the full derivation of the electromagnetic fields in a rectangular waveguide. However, the transverse fields of several of the lowest-order TE_{mn} and TM_{mn} modes are plotted in Figure 4.27. The lowest-order mode, the TE_{10} mode, is a very important mode which we will study in detail next.

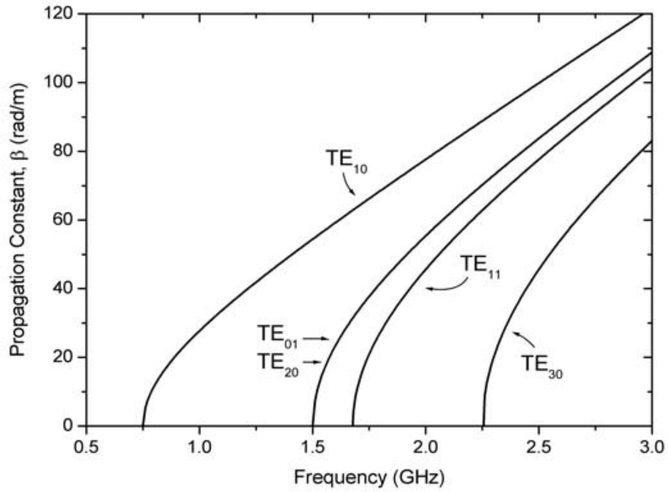


Figure 4.26: Propagation constants of the 5 lowest-order modes of the rectangular metallic waveguide in Example 4.6.

4.5.4 The Fundamental Mode

In this section we focus on the lowest-order mode that can propagate in a rectangular waveguide. Unlike the parallel plate waveguide, there is no TEM mode since the conducting plates force the x- and y-components of the electric field to be zero at the walls of the waveguide. Since the TE_{00}, TM_{00}, TM_{m0}, and TM_{0n} modes do not exist and by definition $a \geq b$, then from Equation 4.102 we see that the mode with the lowest cutoff frequency is the TE_{10} mode. For this mode, the electromagnetic fields are exactly the same as the TE fields of the parallel plate waveguide with $m = 1$:

$$\bar{E} = E_o \sin\left(\frac{\pi}{a}x\right) e^{-j\beta z}\hat{y} \tag{4.104}$$

$$\bar{H} = \frac{E_o}{\omega\mu_o}\left\{-\beta \sin\left(\frac{\pi}{a}x\right)\hat{x} + j\frac{\pi}{a}\cos\left(\frac{\pi}{a}x\right)\hat{z}\right\} e^{-j\beta z} \tag{4.105}$$

From Equation 4.101 the propagation constant of the TE_{10} mode is

$$\beta_{10} = \sqrt{\omega^2\mu_o\varepsilon - (\pi/a)^2} \tag{4.106}$$

and from Equation 4.102 the cutoff frequency is

$$f_{c,10} = \frac{1}{2a\sqrt{\mu_o\varepsilon}} = \frac{c}{2a\sqrt{\varepsilon_r}} \tag{4.107}$$

Notice that a finite electric field exists on the conductors at $y = 0$ and $y = b$. From Equation 4.36, we see that this will give rise to the surface charge densities:

$$\rho_s = \pm\hat{y}\cdot\varepsilon\bar{E} = \pm\varepsilon E_o \sin\left(\frac{\pi}{a}x\right) e^{-j\beta z} \tag{4.108}$$

on the upper and lower walls where the positive and negative signs refer to the walls at $y = 0$ and $y = b$, respectively (see Figure 4.25). In addition, the presence of a nonzero tangential magnetic field will give rise

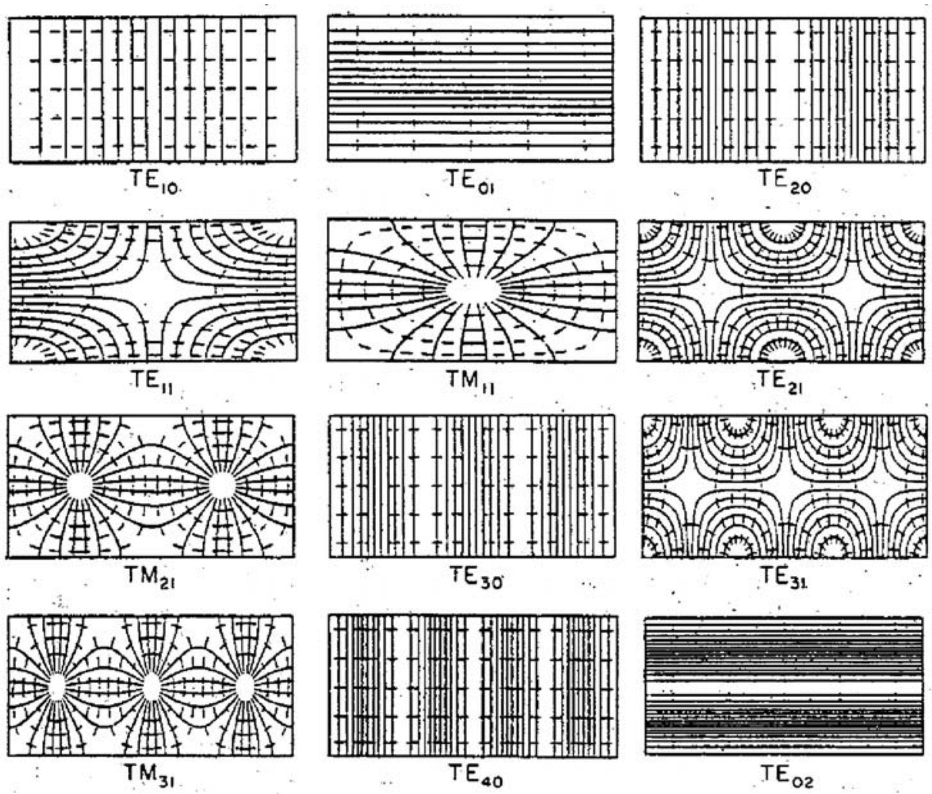


Figure 4.27: Transverse field profiles of the twelve lowest-order TE and TM modes in a rectangular waveguide. Solid lines indicate the electric field, while dashed lines indicate the magnetic field. (Reproduced from C. S. Lee, S. W. Lee and S. L. Chuang. Plot of modal field distributions in rectangular and circular waveguides, *IEEE Trans. Microwave Theory and Techniques*, 33(3), pp. 271-274, March 1985)

to a surface currents on all of the walls, which can be found from Equation 4.37. On the left and right walls, the surface currents are

$$\bar{J}_s = \pm\hat{x} \times \bar{H} = \pm\hat{y}H_z = \pm\hat{y}\frac{jE_o}{\omega\mu_o}\frac{\pi}{a}\cos\left(\frac{\pi}{a}x\right)e^{-j\beta z} \tag{4.109}$$

where the positive and negative signs refer to the walls at $x = 0$ and $x = a$, respectively. On the top and bottom walls, the surface currents are

$$\begin{aligned}\bar{J}_s &= \pm\hat{y} \times \bar{H} = \pm\left(\hat{x}H_z - \hat{z}H_x\right) \\ &= \pm\left[j\frac{\pi}{a}\cos\left(\frac{\pi}{a}x\right)\hat{x} + \beta\sin\left(\frac{\pi}{a}x\right)\hat{z}\right]\frac{E_o}{\omega\mu_o}e^{-j\beta z}\end{aligned} \tag{4.110}$$

where the positive and negative signs refer to the walls at $y = 0$ and $y = b$, respectively.

The time averaged power per unit area flowing down the waveguide in the TE_{10} mode is calculated from the Poynting vector in the z-direction:

$$\left\langle\bar{\mathcal{S}}\right\rangle \cdot \hat{z} = -\frac{1}{2}\mathcal{R}e\left\{E_y H_x^*\right\} \tag{4.111}$$

Substituting the field solutions, this yields

$$\left\langle\bar{\mathcal{S}}\right\rangle \cdot \hat{z} = \frac{\beta|E_o|^2}{2\omega\mu_o}\sin^2\left(\frac{\pi}{a}x\right) \tag{4.112}$$

The total average transmitted power transmitted by the TE_{10} mode in the z-direction is then

$$P = \int_0^b \int_0^a \left(\left\langle \bar{\mathcal{S}} \right\rangle \cdot \hat{z} \right) dx dy = \frac{\beta |E_o|^2}{2\omega\mu_o} \frac{ab}{2} \tag{4.113}$$

Notice that the larger the product ab is, the more power can be transmitted by the waveguide.

4.5.5 Why does $a = 2b$ in a Typical Waveguide?

In rectangular waveguide, we typically want to optimize the power carrying capacity as well as the frequency range in which only the fundamental TE_{10} mode can propagate, i.e the range in which the waveguide is "single-mode" (the reason for this will become apparent in Chapter 5 when we discuss dispersion).

For a rectangular waveguide with dimensions a and b, the mode with the lowest cutoff frequency is the TE_{10} mode. The mode with the next higher cutoff frequency is either the TE_{01} mode or the TE_{20} mode whose cutoff frequencies are found from Equation 4.102 to be:

$$f_{c,01} = \frac{1}{2b\sqrt{\mu_o \varepsilon}} \tag{4.114}$$

$$f_{c,20} = \frac{1}{a\sqrt{\mu_o \varepsilon}} \tag{4.115}$$

Which mode has the next higher cutoff frequency depends upon the relationship between a and b. If $b = a/2$, then $f_{c,01}$ and $f_{c,20}$ are equal. If $a/2 < b < a$, then $f_{c,01} < f_{c,20}$ and the mode with the next higher cutoff frequency is the TE_{01} mode. In this case, the range of frequencies over which single-mode operation occurs is

$$\frac{1}{2a\sqrt{\mu_o \varepsilon}} < f < \frac{1}{2b\sqrt{\mu_o \varepsilon}} \tag{4.116}$$

This range of frequencies is the largest when a and b are chosen to be as far apart as possible. Therefore, we choose $b = a/2$.

If we choose $b < a/2$, then $f_{c,01} > f_{c,20}$ and the mode with the next higher cutoff frequency is the TE_{20} mode. Therefore, the range of frequencies over which single-mode operation occurs is

$$\frac{1}{2a\sqrt{\mu_o \varepsilon}} < f < \frac{1}{a\sqrt{\mu_o \varepsilon}} \tag{4.117}$$

and therefore as long as $b < a/2$ the single-mode frequency range does not depend on b. From Equation 4.113 we know that the larger the product ab is, the more power can be transmitted so we choose the largest possible value for b, which is $b = a/2$. Therefore, to optimize the single-mode frequency range *and* the power handling capability we often find that $a = 2b$ in practical rectangular waveguides.

4.6 Important Concepts

Upon the completion of this chapter, you should be able to:

- Contrast the behavior of waves propagating in unbounded media and in waveguides.
- Derive the general waveguiding condition for a three-layered structure, and how it reduces to the waveguiding conditions for parallel plate metallic waveguides and dielectric planar waveguides.
- For parallel plate waveguides:
  - Determine the propagation constants and cutoff conditions.
  - Describe the behavior of the fields of the TE, TM, and TEM modes.

- Show why the $m = 0$ TE mode cannot exist.
- Derive the currents that are generated in the conducting regions.
- Understand and be able to calculate the "wavelength inside the waveguide," the frequency of operation, the wavelength of operation, the wavelength in the transverse direction inside the waveguide, the phase velocity of the modes, and the guide impedance. You should also know how these parameters depend on the properties of the dielectric between the two plates.
- Calculate the reflection coefficient between two dissimilar waveguides.

- For symmetric dielectric planar waveguides:
  - Determine the propagation constants and cutoff conditions.
  - Describe the behavior of the fields of the TE and TM modes, and understand why no TEM mode exists.
  - Understand the behavior of the fields as a mode approaches or "moves away" from cutoff.
- For rectangular waveguides:
  - Determine the propagation constants and cutoff conditions.
  - Describe in general the behavior of the fields of the TE and TM modes.
  - Understand in detail the lowest order mode and the currents generated in the walls of the waveguide.
  - Understand and be able to calculate the "wavelength inside the waveguide", the frequency of operation, the wavelength of operation, the wavelength in the transverse direction inside the waveguide, and the phase velocity of the modes.
  - Explain why $a = 2b$ in many rectangular waveguides.

Chapter 5

Group Velocity and Dispersion

Up to this point, we have only considered the propagation of monochromatic signals (electromagnetic fields at a single wavelength) that exist over all time. In information systems, we want to send modulated signals or electromagnetic pulses through a system. From Fourier analysis, it can be shown that any time-limited signal or modulated signal actually consists of a spectrum of frequencies. Most often, the different wavelength (frequency) components of an electromagnetic pulse each propagate at a different velocity and therefore become spread out in time as they propagate through a material or waveguide. This is known as dispersion.

5.1 Dispersion of Electromagnetic Signals

To understand how dispersion causes electromagnetic pulses to broaden, consider the simple interpretation shown in Figure 5.1. At $t = t_1$, a pulse consisting of multiple wavelengths is launched into a material or waveguide exhibiting dispersion. Assume for now that all media are lossless. At this point, all of the wavelength components overlap in time. However, as they propagate the different wavelength components travel at different group velocities. At $t = t_2$, the various wavelengths no longer overlap, and the combined pulse shape has broadened in time. It is important to realize that this effect is *not due to attenuation*, since we have assumed that all media are lossless. Instead, it is only due to the "spreading out" of the electromagnetic energy as the pulse propagates. The total energy in the pulse remains the same.

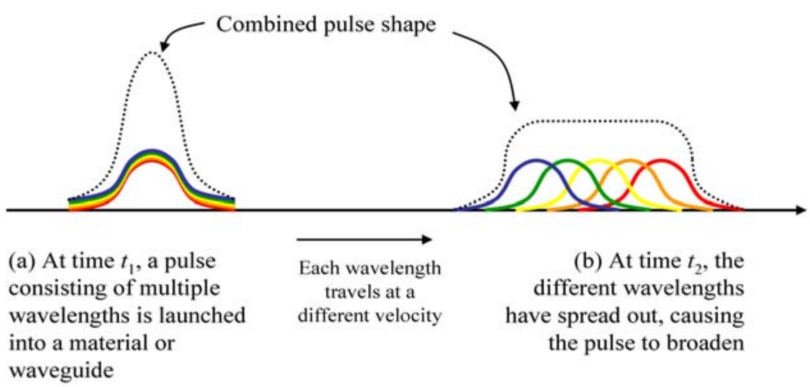


Figure 5.1: Pulse broadening due to dispersion. (a) The initial pulse consists of different wavelengths. Each wavelength propagates at a different velocity due to dispersion. (b) The velocity difference causes the different wavelengths to spread out and the pulse to broaden.

In order to understand propagation in a dispersive system, we first need to understand the difference between "phase velocity" and "group velocity."

5.1.1 Phase Velocity vs. Group Velocity

Up to this point, we have been studying only the propagation of monochromatic electromagnetic waves, which we have assumed exist for all time and all space. We have not considered time-limited signals. What we have been calling "velocity" of the wave is actually the "phase velocity" of the wave. For example, c is the phase velocity of a monochromatic plane wave in free space.

Consider a propagating sinusoidal wave, as illustrated in Figure 5.2. The *phase velocity*, v_p, is the velocity at which a particular point on this wave travels through space. To determine an expression for the phase velocity of the wave, we take a fixed reference point on the wave, and analyze how it moves as time progresses. We will assume the wave is propagating in the +z-direction with a propagation constant, β, and a frequency, ω. To fix a point on the wave, we set $\sin(\omega t - \beta z)$ equal to a constant. Then, to determine how this fixed

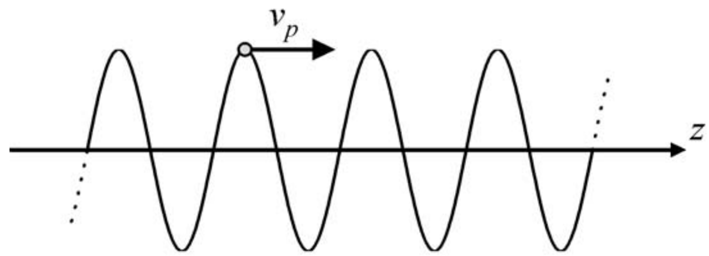


Figure 5.2: Illustration of the phase velocity (v_p) of a wave. The phase velocity is the velocity at which a particular point on the wave travels.

point moves in time, we solve for z which yields:

$$\omega t - \beta z = A \quad \rightarrow \quad z = \frac{\omega}{\beta}t - \frac{A}{\beta} \tag{5.1}$$

where A is an arbitrary constant. The phase velocity of the wave is then determined by the rate at which this point moves in the z-direction:

$$v_p = \frac{dz}{dt} = \frac{\omega}{\beta} \tag{5.2}$$

Equation 5.2 can also be rewritten in terms of the *free space wavelength* of the signal, λ_o, since in free space $\omega = 2\pi c/\lambda_o$:

$$v_p = \frac{2\pi c}{\lambda_o \beta} \tag{5.3}$$

While we have derived the expression for v_p by considering only a propagating wave, the result also applies to a waveguide mode. In this case, we can rewrite the phase velocity in terms of the effective index, n_{eff}, of the guided mode (recall that $k_o = \omega/c$):

$$v_p = \frac{\omega}{\beta} = \frac{k_o c}{\beta} = \frac{c}{n_{\text{eff}}} \tag{5.4}$$

This relationship shows that the phase velocity of a guided mode is inversely proportional to the effective index of that mode. Therefore, as a waveguide mode approaches cutoff, its phase velocity increases.

For a non-monochromatic signal (such as an electromagnetic pulse), we can define another velocity, the *group velocity*, v_g. When a pulse of light is launched into a material or a waveguide, the group velocity is the velocity at which the *pulse envelope* travels. This is illustrated in Figure 5.3. In the figure, the envelope of the pulse is indicated by the dashed lines. The group velocity is also the velocity at which the pulse energy travels.

To determine an expression for the group velocity, we will consider a simple signal consisting of two equal amplitude waves at slightly different frequencies: $\omega_o + \Delta\omega$ and $\omega_o - \Delta\omega$, such that the frequency difference between the waves is $2\Delta\omega$. The propagation constants of these two waves are $\beta_o + \Delta\beta$ and $\beta_o - \Delta\beta$. In time-domain form, the two waves can be written as

$$\mathcal{E}_1 = E_o \cos[(\omega_o + \Delta\omega)\, t - (\beta_o + \Delta\beta)\, z] \quad \text{and} \quad \mathcal{E}_2 = E_o \cos[(\omega_o - \Delta\omega)\, t - (\beta_o - \Delta\beta)\, z] \tag{5.5}$$

To make it easier to add these waves, we write them in the following form:

$$\mathcal{E}_1 = \mathcal{R}e\left\{E_o e^{j[(\omega_o+\Delta\omega)t-(\beta_o+\Delta\beta)z]}\right\} \quad \text{and} \quad \mathcal{E}_1 = \mathcal{R}e\left\{E_o e^{j[(\omega_o-\Delta\omega)t-(\beta_o-\Delta\beta)z]}\right\} \tag{5.6}$$

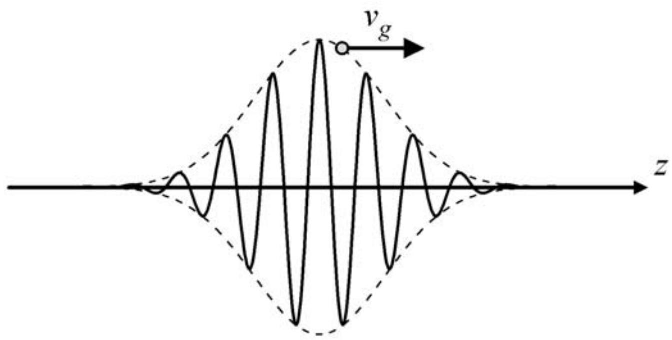


Figure 5.3: Illustration of the group velocity (v_g) of a pulse. The group velocity is the velocity at which the pulse envelope (indicated by dashed lines) travels. It is also the velocity at which the pulse energy travels.

Adding these two waves we obtain

$$\mathcal{E}_{\text{tot}} = \mathcal{E}_1 + \mathcal{E}_2 = \mathcal{R}e\left\{E_o\left[e^{j(\Delta\omega t - \Delta\beta z)} + e^{-j(\Delta\omega t - \Delta\beta z)}\right]e^{j(\omega_o t - \beta_o z)}\right\}$$

$$= \mathcal{R}e\left\{2E_o\cos(\Delta\omega t - \Delta\beta z)e^{j(\omega_o t - \beta_o z)}\right\}$$

$$= 2E_o\cos(\Delta\omega t - \Delta\beta z)\cos(\omega_o t - \beta_o z) \tag{5.7}$$

The two waves form a "beat pattern" as illustrated in Figure 5.3. The envelope of this beat pattern, indicated by the dashed lines in the figure, is $2E_o\cos(\Delta\beta z - \Delta\omega t)$. The group velocity of the beat pattern is the velocity at which this envelope travels.

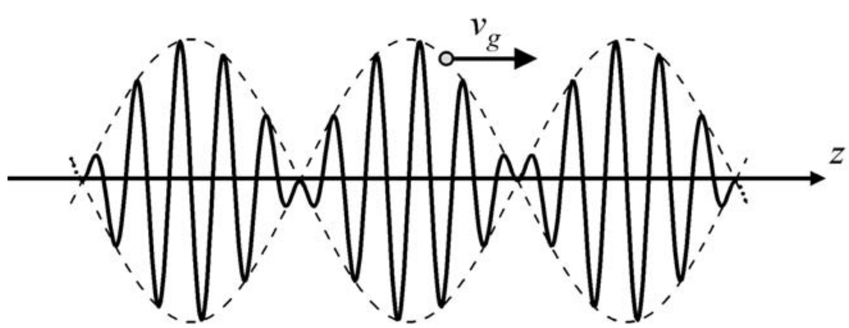


Figure 5.4: Illustration of the "beat pattern" formed by two waves with slightly different frequencies. The group velocity (v_g) is the velocity at which the "beat" envelope travels.

To determine the group velocity, we now take a fixed reference point on the envelope and determine the rate at which it travels. Therefore, we set the *envelope,* $E_o\cos(\Delta\omega t - \Delta\beta z)$, equal to a constant and solve for z. This yields:

$$\Delta\omega t - \Delta\beta z = A \quad \rightarrow \quad z = \frac{\Delta\omega}{\Delta\beta}t - \frac{A}{\Delta\beta} \tag{5.8}$$

where A is an arbitrary constant. The group velocity of the wave is then determined by the rate at which this point moves in the z-direction:

$$\frac{dz}{dt} = \frac{\Delta\omega}{\Delta\beta} \tag{5.9}$$

A realistic electromagnetic signal does not consist only of two frequencies, but of a continuum of frequencies around a central frequency, ω_o. To find the group velocity at ω_o, we find the limit of Equation 5.9 as $\Delta\omega$ and $\Delta\beta$ approach zero:

$$v_g = \lim_{\Delta\omega \to 0} \frac{\Delta\omega}{\Delta\beta} = \frac{d\omega}{d\beta} \tag{5.10}$$

The group velocity can also be expressed in terms of the *free space wavelength* of the signal, λ_o, since in free space $\omega = 2\pi c/\lambda_o$. Using the chain rule we find

$$v_g = \frac{d\omega}{d\lambda_o}\frac{d\lambda_o}{d\beta} = \frac{d}{d\lambda_o}\left(\frac{2\pi c}{\lambda_o}\right)\frac{d\lambda_o}{d\beta} = \frac{-2\pi c}{\lambda_o^2}\frac{d\lambda_o}{d\beta} \tag{5.11}$$

An important point to note about the group velocity is that if β is a *linear function* of ω, then the phase velocity and the group velocity are equal. For example, in free space the propagation constant of an electromagnetic wave is $k = \omega\sqrt{\mu_o\varepsilon_o} = \omega/c$, which is a linear function of frequency. From Equation 5.2, the phase velocity in free space is

$$v_p = \frac{\omega}{k} = \frac{kc}{k} = c \tag{5.12}$$

as expected. The group velocity of the wave in free space is

$$v_g = \frac{d\omega}{dk} = \frac{d}{dk}(kc) = c \tag{5.13}$$

Notice that $v_p = v_g$ since the propagation constant is a linear function of frequency! It is only when β is a *nonlinear function* of ω that the phase velocity and the group velocity are different!

Once the group velocity of a pulse is known, we can calculate the *group delay*, τ_g, which is the time required for a particular frequency component of the pulse to propagate a distance *L:*

$$\tau_g = \frac{L}{v_g} = L\frac{d\beta}{d\omega} = \frac{-\lambda_o^2 L}{2\pi c}\frac{d\beta}{d\lambda_o} \tag{5.14}$$

From Equation 5.14 it is apparent that the group delay is a function of frequency (or wavelength) provided that β is a nonlinear function of ω. This means that each frequency/wavelength component of a pulse will arrive at a different time, causing what is known as *dispersion.* It is important to notice that dispersion is *not* caused by different waveguide modes propagating at different velocities! (This is a separate effect, called "multimode distortion," which we will not discuss here). Dispersion occurs even when all of the energy in a waveguide is confined to one waveguide mode.

The type of dispersion a pulse experiences is classified according to the relative magnitudes of the phase and group velocities, which are determined by the curvature of the ω vs. β graph. As stated previously, if the propagation constant is a linear function of frequency as shown in Figure 5.5(a), then $v_p = v_g$ and a pulse will not experience dispersion. This situation is called *dispersion-less.* If the ω vs. β graph exhibits a "downward" curvature as shown in Figure 5.5(b), then $v_p > v_g$ and the pulse experiences *normal dispersion.* It is referred to as normal because in a material such as glass the index of refraction increases towards larger frequencies, which decreases the linear increase of ω vs. β curve, which then leads to $v_p < v_g$. Finally, if the ω vs. β graph has an "upwards" curvature as shown in Figure 5.5(c), then $v_p > v_g$ and the pulse experiences *anomalous dispersion.*

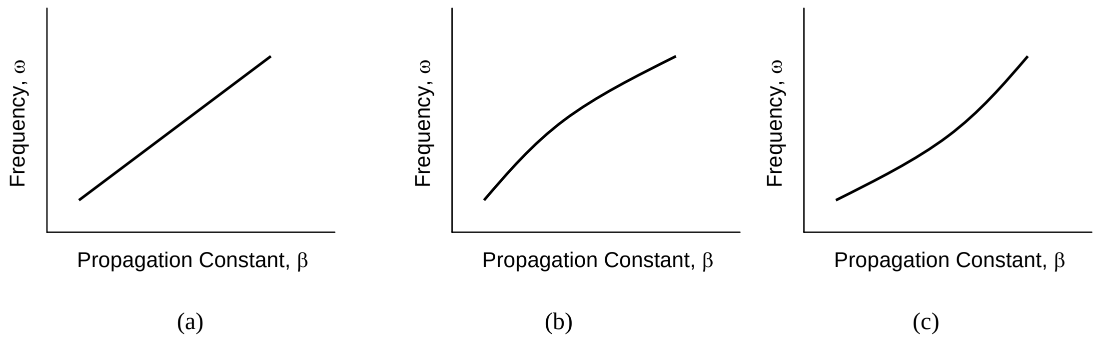


Figure 5.5: The dispersion type can be determined by the slope of the ω vs. β graph. (a) Dispersion-less ($v_p = v_g$). (b) Normal dispersion ($v_p < v_g$). (c) Anomalous dispersion ($v_p > v_g$).

5.1.2 The Origins of Dispersion

What are the causes of dispersion? Or in other words, what causes β to be a nonlinear function of ω? In general, dispersion is the result of two effects:

1. **Material dispersion:** Dispersion caused by the material through which the pulse is propagating. The index of refraction (or permittivity) of most materials varies with frequency, causing the propagation constant to be a nonlinear function of ω.

2. **Waveguide dispersion:** Dispersion caused by confining the electromagnetic energy within a waveguide. As we saw in Chapter 4, the geometry of a waveguide itself causes β to be a nonlinear function of ω, even if we assume that the material parameters are constant.

In a waveguide, both types of dispersion generally exist. However, in many situations material dispersion is much smaller than waveguide dispersion and can be neglected. One example of when this is *not* the case is in optical fibers. For fibers, material and waveguide dispersion are similar in magnitude and both must be considered. In the following sections, we will investigate the material dispersion properties of a plasma, and the waveguide dispersion present in a parallel plate waveguide.

5.2 Material Dispersion: Plasma

In Chapter 2, we discussed a plasma, which is a collection of oscillating positively and negatively charged particles in which the time-averaged charge density is zero. Below the plasma frequency, electromagnetic waves propagate without loss; however, the plasma also exhibits material dispersion. This can be demonstrated as follows. In Chapter 2, we found that the permittivity of a plasma is given by

$$\varepsilon_p = \varepsilon_o \left(1 - \frac{\omega_p^2}{\omega^2}\right) \tag{5.15}$$

where ω_p is the plasma frequency. Note that the permittivity is highly dependent upon frequency. The propagation constant of an electromagnetic wave in a plasma is

$$k_p = \omega\sqrt{\mu_o\varepsilon_o\left(1 - \frac{\omega_p^2}{\omega^2}\right)} = k_o\sqrt{1 - \frac{\omega_p^2}{\omega^2}} = \frac{1}{c}\sqrt{\omega^2 - \omega_p^2} \tag{5.16}$$

If $\omega > \omega_p$, then $\varepsilon_p > 0$ and k_p is real, corresponding to a propagating wave. Solving Equation 5.16 for ω we find

$$\omega = \sqrt{c^2k_p^2 + \omega_p^2} \tag{5.17}$$

The functional dependence of ω on k_p is shown in Figure 5.6. The dotted line corresponds to the propagation constant in free space, $k_o = \omega\sqrt{\mu_o\varepsilon_o}$. Notice that the propagation constant is a nonlinear function of frequency, which means that dispersion will occur when a signal propagates through a plasma.

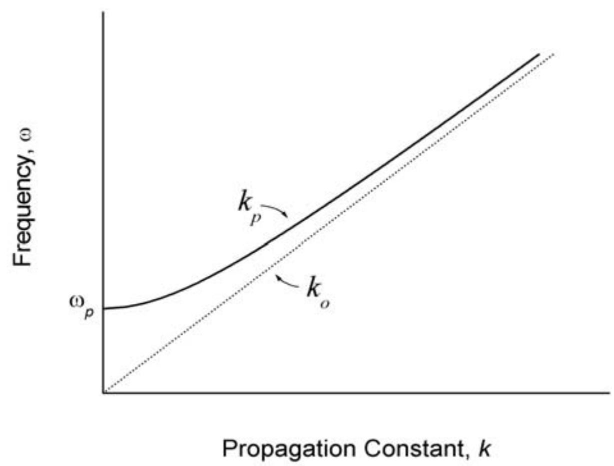


Figure 5.6: ω-k relationship for a wave in a plasma.

To find the group velocity, we could solve for ω, then find $d\omega/dk_p$. However, it is often easier to calculate the inverse as follows:

$$v_g = \frac{d\omega}{dk_p} = \left(\frac{dk_p}{d\omega}\right)^{-1} \tag{5.18}$$

This is mathematically correct provided that the curve of k_p versus ω is continuous and single-valued, which holds for the cases of interest to us. Using this method, the group velocity of a pulse propagating through a plasma is

$$v_g = \frac{d\omega}{dk_p} = \left(\frac{dk_p}{d\omega}\right)^{-1} = \left(\frac{\omega}{c\sqrt{\omega^2 - \omega_p^2}}\right)^{-1} = c\sqrt{1 - \frac{\omega_p^2}{\omega^2}} \tag{5.19}$$

(Refer to Appendix E for the alternate method of derivation.) The group velocity as a function of frequency is shown in Figure 5.7(a). Notice that v_g is a nonlinear function that is highly dependent upon frequency, and decreases towards zero near the plasma frequency, ω_p.

The group delay over a distance L through the plasma (where $\omega > \omega_p$) is

$$\tau_g = \frac{L}{v_g} = \frac{L}{c\sqrt{1 - \omega_p^2/\omega^2}} \tag{5.20}$$

The group delay per unit length (τ_g/L) is shown in Figure 5.7(b). As the frequency approaches the plasma frequency, the propagation time increases dramatically. Notice also that the propagation time approaches that of a wave in free space as $\omega \to \infty$. The material dispersion is clearly evident in the figure.

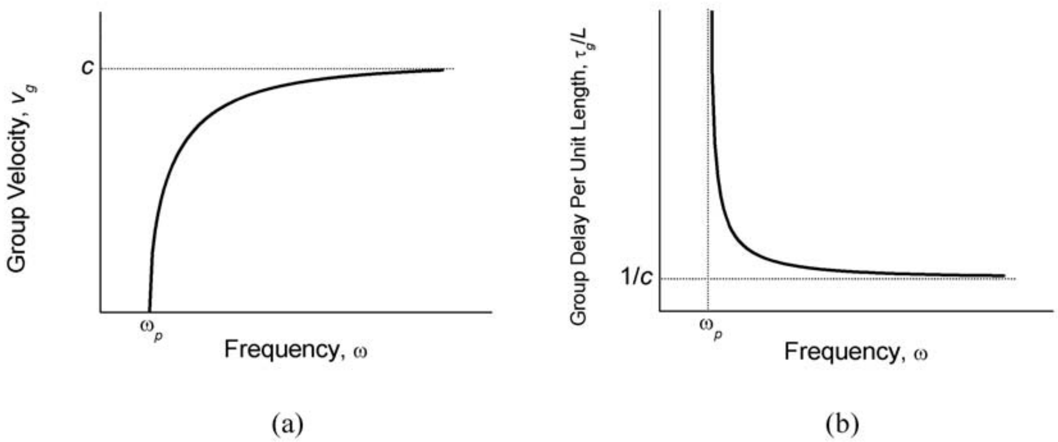


Figure 5.7: (a) Group velocity versus frequency for a plasma. (b) Group delay per unit length in a plasma. The propagation time increases dramatically as the frequency approaches the plasma frequency.

$\lll\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 5.1:

As stated previously, the index of refraction of most materials varies with wavelength . Suppose you have a fictitious glass whose index of refraction can be expressed by the following function over the wavelength range of interest:

$$n(\lambda_o) = A + \frac{B}{\lambda_o} + \frac{C}{\lambda_o^2}$$

where λ_o is the free space wavelength, and A, B, and C are constants. Find expressions for the phase and group velocities in terms of the wavelength and the constants A, B, and C.

Solution:

We must first obtain an expression for the propagation constant of a wave in this glass as a function of wavelength:

$$\beta(\lambda_o) = n(\lambda_o)k_o = n(\lambda_o)\frac{2\pi}{\lambda_o} = 2\pi(A\lambda_o^{-1} + B\lambda_o^{-2} + C\lambda_o^{-3})$$

From Equation 5.3, the phase velocity is

$$v_p = \frac{2\pi c}{\lambda_o \beta} = \frac{2\pi c}{\lambda_o 2\pi(A\lambda_o^{-1} + B\lambda_o^{-2} + C\lambda_o^{-3})} = \frac{c}{A + B\lambda_o^{-1} + C\lambda_o^{-2}}$$

From Equation 5.11, the group velocity is

$$v_g = \frac{-2\pi c}{\lambda_o^2}\frac{d\lambda_o}{d\beta} = \frac{-2\pi c}{\lambda_o^2}\left(\frac{d\beta}{d\lambda_o}\right)^{-1} = \frac{-2\pi c}{\lambda_o^2}\left\{2\pi(-A\lambda_o^{-2} - 2B\lambda_o^{-3} - 3C\lambda_o^{-4})\right\}^{-1}$$

$$= \frac{c}{A + 2B\lambda_o^{-1} + 3C\lambda_o^{-2}}$$

The same result for v_g would have been obtained if we would have found an expression for λ_o in terms of β and then directly solved for the derivative $d\lambda_o/d\beta$. This is a daunting task, however (try it for yourself!). Notice that if the index of refraction is a constant (B and C are zero) then the phase and group velocities are equal.

$\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg$

5.3 Waveguide Dispersion: Parallel Plate Waveguide

In order to understand waveguide dispersion, we will focus on the dispersion properties of the simplest waveguide we have studied so far: the parallel plate waveguide. We will solve for the waveguide dispersion in the absence of material dispersion, i.e. assuming that the material properties are not functions of frequency. Using the expression for the propagation constant we have previously derived (see Chapter 4), we solve for ω to obtain an expression for the frequency in terms of β for the m^{th} mode:

$$\omega = \frac{c}{\sqrt{\varepsilon_r}}\sqrt{\beta_m^2 + (m\pi/a)^2} \tag{5.21}$$

Assuming that ε_r is not a function of frequency (i.e. there is no material dispersion), the group velocity of the m^{th} mode is then

$$v_{g,m} = \frac{d\omega}{d\beta_m} = \frac{c\beta_m}{\varepsilon_r k_o} = \frac{c}{\sqrt{\varepsilon_r}}\sqrt{1 - f_{c,m}^2/f^2} \tag{5.22}$$

The details of this derivation are given in Appendix E. The group velocities of the five lowest-order TM modes of a parallel plate waveguide are shown in Figure 5.8(a). Notice that the group velocity of the TEM ($m = 0$) mode is always $v_{g,0} = c/\sqrt{\varepsilon_r}$ regardless of frequency! In fact, this is equal to the group velocity of a plane wave propagating in a dielectric with a relative permittivity ε_r. This makes intuitive sense, as we previously found that the TEM mode resembles a plane wave that is truncated at the perfectly conducting top and bottom plates. This is one reason why the TEM mode is extremely important!

From Equation 5.14 we find the group delay, τ_g, for each mode:

$$\tau_g = \frac{L}{v_g} = \frac{L\sqrt{\varepsilon_r}}{c\sqrt{1 - f_{c,m}^2/f^2}} \tag{5.23}$$

The group delay per unit length (τ_g/L) in a parallel plate waveguide is shown in Figure 5.8(b). Notice that the TEM mode has a constant group delay regardless of frequency. For the remainder of the modes, however, the propagation time is highly dependent upon frequency, which leads to dispersion.

$\lll\lll\lll\lll\lll\ggg\ggg\ggg\ggg\ggg$

EXAMPLE 5.2:

Consider an air-filled ($\varepsilon_r = 1$) parallel plate waveguide with a plate spacing of 5 cm. What is the group velocity of a narrowband signal with a center frequency of 10 GHz, propagating in the lowest-order TE mode? How long will this signal require to propagate 10 m? How long would a pulse at 5 GHz require to propagate the same distance?

Solution:

The lowest-order TE mode is the TE_1 mode which has a cutoff frequency of

$$f_{c,1} = \frac{c}{2a\sqrt{\varepsilon_r}} = \frac{3\times10^8}{2\,(0.05)} = 3\,\text{GHz}$$

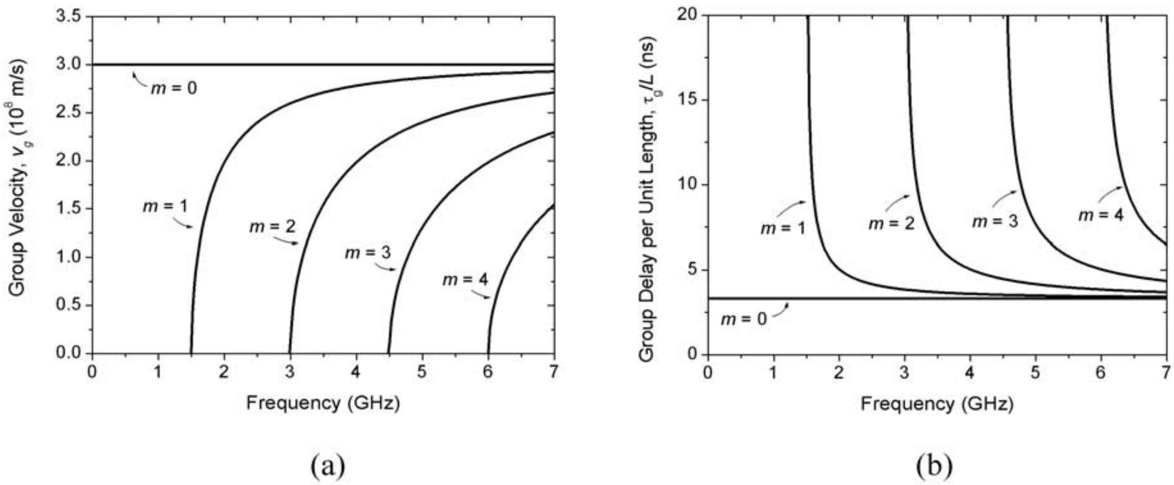


Figure 5.8: (a) Group velocity (v_g) of the four lowest-order TM modes of a parallel plate waveguide. Notice that the TEM ($m = 0$) mode has a constant velocity regardless of frequency, and therefore exhibits no group velocity dispersion. (b) Group delay per unit length (τ_g/L) of the four lowest-order TM modes of a parallel plate waveguide. (Parameters: $a = 10$ cm, $\varepsilon_r = 1$)

From Equation 5.22, the group velocity is then

$$v_{g,1} = \frac{c}{\sqrt{\varepsilon_r}}\sqrt{1 - f_{c,m}^2/f^2}$$

$$= \left(3 \times 10^8\right)\sqrt{1 - 3^2/10^2} = 2.86 \times 10^8 \quad \text{m/s}$$

The time required to propagate 10 m (i.e. the group delay) is

$$\tau_g = \frac{L}{v_g} = \frac{10}{2.86 \times 10^8} = 34.9\,\text{ns}$$

Using the same method, we find the group velocity at 5 GHz is 2.40×10^8 m/s. The group delay at 5 GHz is then

$$\tau_g = \frac{10}{2.40 \times 10^8} = 41.7\,\text{ns}$$

Therefore, a pulse at 10 GHz would arrive almost 7 ns sooner than the pulse at 5 GHz, which would lead to dispersion in a pulse containing both of these frequencies.

$\lll\lll\lll\lll\ggg\ggg\ggg\ggg$

5.4 Important Concepts

Upon the completion of this chapter, you should be able to:

- Describe in detail the importance of group velocity dispersion in data communication systems and its relation to pulse distortion.

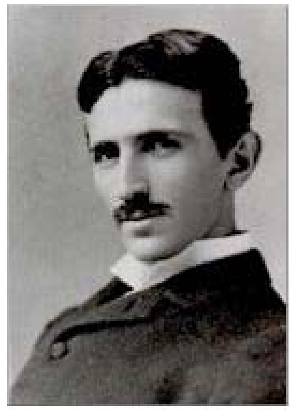
- Derive expressions for the phase and group velocity of a electromagnetic signal in terms of the frequency and propagation constant.
- Understand the difference between material and waveguide dispersion, and the origins of both types of dispersion.
- Derive the group velocity and group delay given an expression for the propagation constant in terms of frequency.
- Understand why the TEM mode does not exhibit waveguide dispersion.

Chapter 6

Radiation

Guglielmo Marconi (1874 - 1937)

"Heinrich Hertz's classical experiments were conducted in his laboratory using a small end-loaded dipole driven by an induction coil and a spark gap for his transmitter. His receiver was a small loop, and detection was by induced sparking. Since the frequency generated by a spark transmitter is determined by the resonant mean frequency of the antenna system, his experiments in 1887 were conducted at VHF/UHF (60 to 500 Mhz) – the corresponding wavelengths (5.0 to 0.6 metre) being practical for indoor experiments. Marconi started experimenting with Hertz's apparatus in 1894. He was fascinated by the idea that by means of Hertzian waves it might be possible to send telegraph signals, without wires, far enough for such a system to have commercial value. By 1896 he achieved a transmission distance of 2.5 kilometres, by using an earth and an elevated aerial at both transmitter and receiver (nowadays called a Marconi antenna). His first permanent station established a link between the Isle of Wight and Bournemouth, England, some 22 km away (in 1897). He established communications across the Channel in 1899. By now he must have been using frequencies in the low HF band, since his aerial systems were much larger. In 1900 he decided to try and achieve transatlantic communications. The required aerial size, and so the signalling frequency, at best could only be projected by extrapolation from values successful over a range of much shorter distances. The aerial at Poldhu, Cornwall in December 1901.... radiated signals in the MF band (about 850 kHz)" [8]



Nikola Tesla (1856 - 1943)

In 1893 the Chicago World Columbian Exposition was lighted by means of Tesla's system and work was begun on the installation of power machinery at Niagara Falls. In a lecture-demonstration given in St. Louis in the same year - two years before Marconi's first experiments - Tesla also predicted wireless communication; the apparatus that he employed contained all the elements of spark and continuous wave that were incorporated into radio transmitters before the advent of the vacuum tube. Engrossed as he was with the transmission of substantial amounts of power, however, he almost perversely rejected the notion of transmission by Hertzian waves, which he considered to be wasteful of energy. He thus proposed wireless communication by actual conduction of electricity through natural media, and, working in Colorado Springs, Colorado, in 1899-1900, proved the earth to be a conductor ... Of his wireless system, he wrote in 1900: 'I have no doubt that it will prove very efficient in enlightening the masses, particularly in still uncivilized countries and less accessible regions, and that it will add materially to general safety, comfort and convenience, and maintenance of peaceful relations.' With the financial backing of J. P. Morgan, he began work on a worldwide communications system, and a 200-foot transmission tower was constructed at Shoreham, on Long Island. By 1905, however, Morgan had withdrawn his support, and the project came to an end. The tower was destroyed by dynamite, under mysterious circumstances, in 1914." [9]

In our study up to this point we have simply assumed that electromagnetic waves exist, and we looked at the propagation of those waves in various media and waveguides. We now turn to the question: how are electromagnetic waves created? The answer, as we will see, is that they are created by radiation from time-varying currents and charges. In this chapter, we will look at radiation from such sources in free space. The study of electromagnetic radiation also forms the bridge between electrostatics and the propagation of electromagnetic waves. Calculations involving radiation require advanced vector mathematics; therefore, it is important that you review the vector theorems given in Chapter 1 to understand the material in this chapter. In addition, it is important to be very familiar with the spherical coordinate system including the unit vectors, discussed in Chapter 1 and also repeated in Figure 6.1 for convenience.

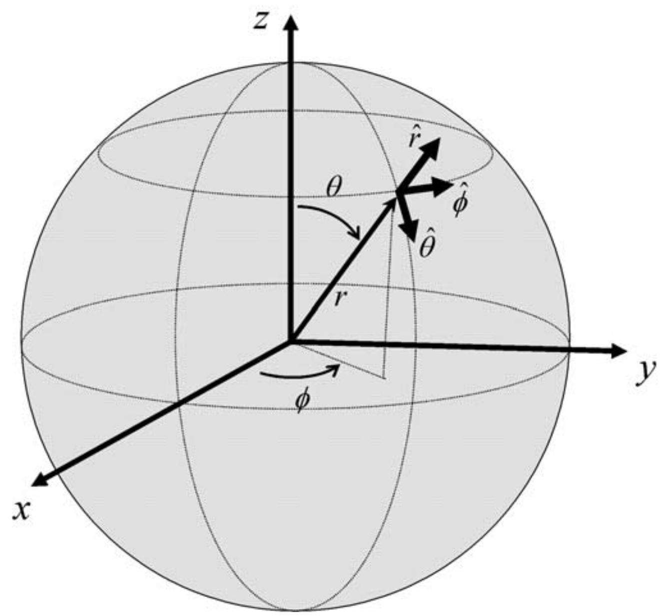


Figure 6.1: The spherical coordinate system and definition of the unit vectors.

6.1 Electrostatics Review

Before investigating the generation of electromagnetic waves, let us first review electrostatics, i.e. when the electromagnetic field quantities do not depend on time. In this case, all of the derivatives with respect to time are equal to zero. Faraday's law in the electrostatic case is:

$$\nabla \times \bar{\mathcal{E}}(\bar{r}) = 0 \tag{6.1}$$

Ampere's law in the electrostatic case is:

$$\nabla \times \bar{\mathcal{H}}(\bar{r}) = \bar{\mathcal{J}}(\bar{r}) \tag{6.2}$$

where $\bar{\mathcal{J}}(\bar{r}) = \sigma\bar{\mathcal{E}}(\bar{r})$ is a direct current source in this context. Gauss' laws in the electrostatic case (in free space) are:

$$\nabla \cdot \bar{\mathcal{E}}(\bar{r}) = \frac{\rho_v(\bar{r})}{\varepsilon_o} \tag{6.3}$$

$$\nabla \cdot \bar{\mathcal{H}}(\bar{r}) = 0 \tag{6.4}$$

Notice that the electric and magnetic fields are independent in the electrostatic case in a lossless medium. This is an important conclusion! Up to this point we have been considering time-varying fields, and a time-varying electric field always gives rise to a time-varying magnetic field, and vice versa. In the electrostatic case, however, we can have an electric field without a magnetic field, or a magnetic field without an electric field!

6.1.1 Scalar Potential

We next review the methods for solving for the electric field when dealing with an electrostatic situation. We will use the scalar potential, which will give us hints about how to solve more complicated problems using the concept of a vector potential. In electrostatics, the solution of the electric field from Equations 6.1 and 6.3 is simplified by exploiting the fact that the curl of the gradient of a scalar field $\Phi(\bar{r})$ is zero, i.e.

$$\nabla \times \left[\nabla \Phi(\bar{r})\right] = 0 \tag{6.5}$$

(refer to Chapter 1) and realizing that the solution of a *scalar* field is easier than a *vector* field. There is some flexibility in how the scalar field (also referred to as the scalar potential) is defined. A convenient choice is

$$\bar{\mathcal{E}}\left(\bar{r}\right) = -\nabla \Phi(\bar{r}) \tag{6.6}$$

where $\Phi(\bar{r})$ is the *scalar potential*, with units of volts. The negative sign is chosen so that electric field lines point from regions of high potential to regions of low potential. Notice how this choice of $\bar{\mathcal{E}}(\bar{r})$ satisfies Equation 6.1. By substituting $\bar{\mathcal{E}}(\bar{r})$ into Gauss' law (Equation 6.3), we obtain the *scalar Poisson's equation*

$$\nabla^2 \Phi = -\frac{\rho_v(\bar{r})}{\varepsilon_o} \tag{6.7}$$

The solution for the scalar potential can be found either from differential or integral methods. The scalar potential for an isolated charge is particularly simple using an integral approach. Starting from Equation 6.3 and using Gauss' theorem (see Chapter 1)

$$\oint_S \left(\bar{A} \cdot \hat{n}\right) dS = \int_V \left(\nabla \cdot \bar{A}\right) dV \tag{6.8}$$

it follows that

$$\oint_S \left(\bar{\mathcal{E}} \cdot \hat{n}\right) dS = \int_V \left(\nabla \cdot \bar{\mathcal{E}}\right) dV = \int_V \left(\frac{\rho_v}{\varepsilon_o}\right) dV = \frac{\text{charge enclosed in volume } V}{\varepsilon_o} \tag{6.9}$$

This is Gauss' Law from electrostatics in integral form.

Let us use this opportunity to review how we utilize Gauss' law to obtain the electric field. For a point charge q at the origin, we expect that the electric field will only have a radial component due to spherical symmetry, and at a fixed radius the electric field will be constant. The field is expressed in spherical coordinates as

$$\bar{\mathcal{E}}(r) = E(r)\hat{r} \tag{6.10}$$

We next choose a spherical Gaussian surface that is concentric with the origin and at a radius r_o, as shown in Figure 6.2. Due to symmetry, we can assume that the field has a constant value over the surface S. Applying Gauss' law (Equation 6.9) we obtain the following, noting that $\hat{n} = \hat{r}$ on the spherical surface:

$$\oint_S \left(\bar{\mathcal{E}} \cdot \hat{n}\right) dS = \oint_S \left[E(r_o)\hat{r} \cdot \hat{r}\right] dS = E(r_o) \oint_S dS = 4\pi r_o^2 E(r_o) = \frac{q}{\varepsilon_o} \tag{6.11}$$

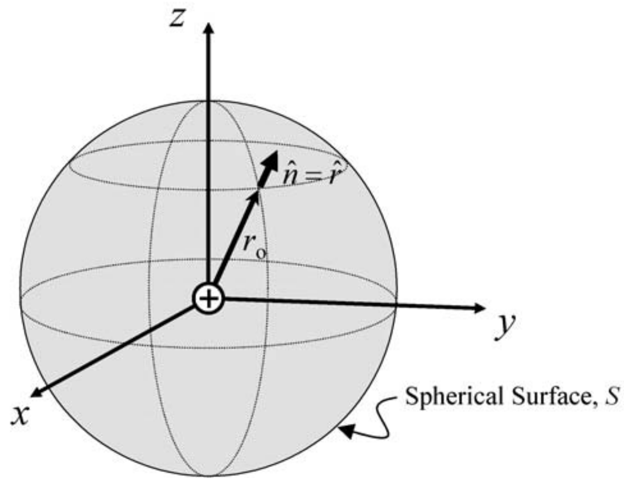


Figure 6.2: A point charge q at the origin. The spherical surface determines the surface of integration to calculate the electric field.

From this it follows that

$$E(r_o) = \frac{q}{4\pi\varepsilon_o r_o^2} \tag{6.12}$$

Even though we have assumed that the surface lies at a particular radius r_o, there is nothing special about this radius since the result in Equation 6.11 will follow for any radius r. Therefore it follows that

$$\bar{\mathcal{E}}(r) = E(r)\hat{r} = \frac{q}{4\pi\varepsilon_o r^2}\hat{r} \tag{6.13}$$

To obtain the scalar potential for the point charge, first note that the gradient in spherical coordinates is (refer to Chapter 1)

$$\nabla\Phi(\bar{r}) = \frac{\partial\Phi(\bar{r})}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial\Phi(\bar{r})}{\partial\theta}\hat{\theta} + \frac{1}{r\sin\theta}\frac{\partial\Phi(\bar{r})}{\partial\phi}\hat{\phi} \tag{6.14}$$

Therefore, from Equation 6.6 the scalar potential for a point charge must satisfy

$$\bar{\mathcal{E}}(r) = \frac{q}{4\pi\varepsilon_o r^2}\hat{r} = -\frac{\partial\Phi(\bar{r})}{\partial r}\hat{r} \tag{6.15}$$

The scalar potential must then be

$$\Phi(r) = \frac{q}{4\pi\varepsilon_o r} + \Phi_o \tag{6.16}$$

where Φ_o is a constant. Imposing the boundary condition that the scalar potential must decrease toward zero as we move away from the charge, i.e. $\Phi(\infty) = 0$, it follows that

$$\Phi = \frac{q}{4\pi\varepsilon_o r} \tag{6.17}$$

If the charge is not at the origin but at the point described by the position vector $\bar{r}'$, then the scalar potential becomes

$$\Phi = \frac{q}{4\pi\varepsilon_o|\bar{r} - \bar{r}'|} \tag{6.18}$$

This solution can be extended to a charge distribution $\rho_v(\bar{r}')$

$$\Phi = \frac{1}{4\pi\varepsilon_o}\int_{V'}\frac{\rho_v(\bar{r}')}{|\bar{r} - \bar{r}'|}dV' \tag{6.19}$$

as shown in Figure 6.3, which is a general integral solution for the scalar Poisson's equation.

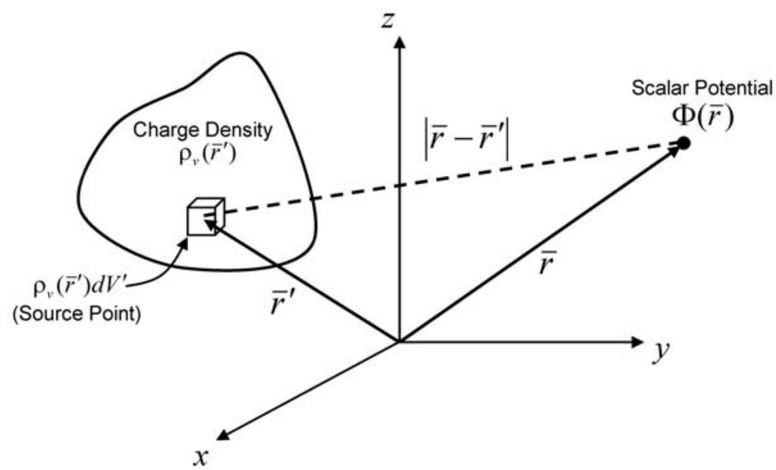


Figure 6.3: Illustration of the scalar potential generated by a charge density located an an arbitrary position in space.

6.1.2 The Electric Dipole

Consider two static charges $+q$ and $-q$ positioned on the z-axis at $z = +d/2$ and $z = -d/2$ respectively, where d is much smaller than the distance to the observation point, as shown in Figure 6.4. What is the electric field far from the origin in this case?

From superposition, the total scalar potential at a radius r is the sum of the scalar potential of the two individual charges:

$$\Phi(r) = \frac{q}{4\pi\varepsilon_o R_+} - \frac{q}{4\pi\varepsilon_o R_-} \tag{6.20}$$

The problem is now focused on obtaining expressions for the distance from the observation point to the positive charge, R_+, and the distance to the negative charge, R_-. In terms of the angle θ, the distance to the positive and negative charges are approximately

$$R_+ \approx r - \frac{d}{2}\cos\theta \tag{6.21}$$

$$R_- \approx r + \frac{d}{2}\cos\theta \tag{6.22}$$

if the observation point is far from the charges (refer to Figure 6.4). We also know that $d \ll r$, and therefore (see Appendix F)

$$\frac{1}{R_+} \approx \frac{1}{r}\left(1 + \frac{d}{2r}\cos\theta\right) \tag{6.23}$$

and

$$\frac{1}{R_-} \approx \frac{1}{r}\left(1 - \frac{d}{2r}\cos\theta\right) \tag{6.24}$$

The total scalar potential is then approximately

$$\Phi(r) \approx \frac{q}{4\pi\varepsilon_o}\frac{1}{r}\left(1 + \frac{d}{2r}\cos\theta\right) - \frac{q}{4\pi\varepsilon_o}\frac{1}{r}\left(1 - \frac{d}{2r}\cos\theta\right) = \frac{qd\cos\theta}{4\pi\varepsilon_o r^2} \tag{6.25}$$

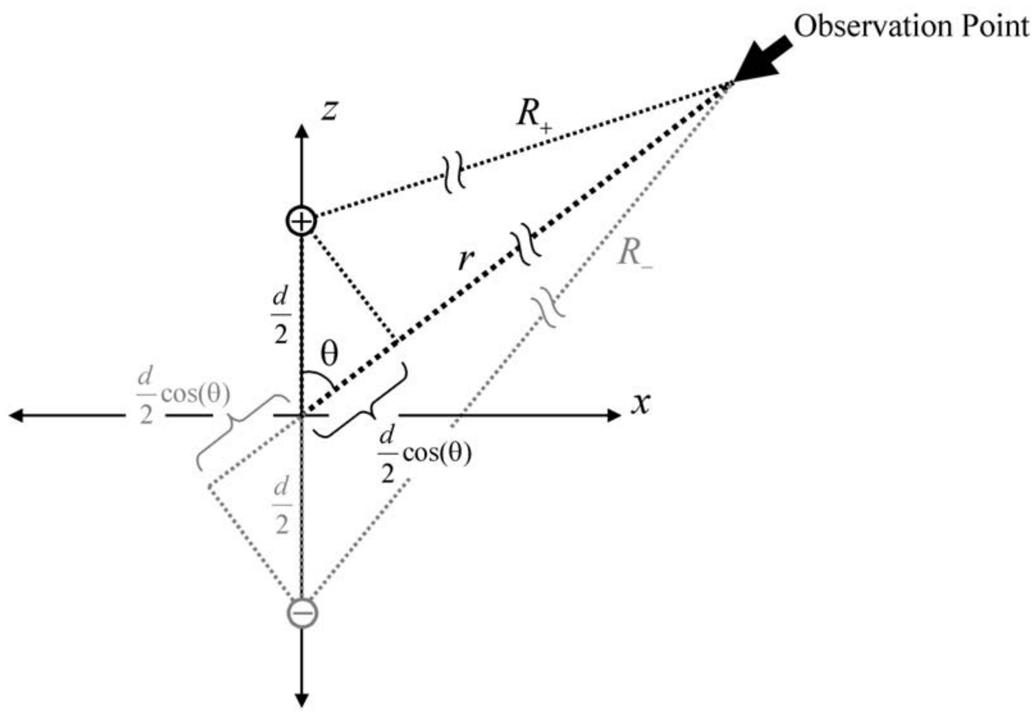


Figure 6.4: Two static charges separated by a distance d. The separation between the charges, d, is assumed to be much less than the distance to the observation point, r.

Solving for the electric field, we obtain

$$\bar{\mathcal{E}}(r,\theta,\phi) = -\nabla\Phi(r) = \frac{qd\cos\theta}{2\pi r^3\varepsilon_o}\hat{r} + \frac{qd\sin\theta}{4\pi r^3\varepsilon_o}\hat{\theta} \tag{6.26}$$

It is customary to define a dipole moment $p = qd$ such that

$$\bar{\mathcal{E}}(r,\theta,\phi) = \frac{p\cos\theta}{2\pi r^3\varepsilon_o}\hat{r} + \frac{p\sin\theta}{4\pi r^3\varepsilon_o}\hat{\theta} \tag{6.27}$$

6.1.3 Vector Potential

If a scalar potential can be used to solve for the electric field, can a potential be defined for the solution of the magnetic field? Recall that the divergence of the magnetic field must be zero (Equation 6.4). Since the divergence of the curl of a vector field is zero (see Chapter 1), i.e.

$$\nabla\cdot\left[\nabla\times\bar{\mathcal{A}}(\bar{r})\right] = 0 \tag{6.28}$$

then $\bar{\mathcal{H}}(\bar{r})$ can be described by a curl of some vector field. A convenient choice is

$$\bar{\mathcal{H}}(\bar{r}) = \frac{1}{\mu_o}\nabla\times\bar{\mathcal{A}}\left(\bar{r}\right) \tag{6.29}$$

where $\bar{\mathcal{A}}(\bar{r})$ is the vector potential. Substituting $\bar{\mathcal{H}}(\bar{r})$ into Ampere's law (Equation 6.2) we obtain

$$\nabla\times\left(\nabla\times\bar{\mathcal{A}}\right) = \nabla\left(\nabla\cdot\bar{\mathcal{A}}\right) - \nabla^2\bar{\mathcal{A}} = \mu_o\bar{\mathcal{J}} \tag{6.30}$$

A vector field is uniquely specified, up to an additive constant, when both the curl and divergence of the vector field are known. In the specification of the divergence of $\bar{\mathcal{A}}$ there is again some flexibility and one convenient choice is

$$\nabla \cdot \bar{\mathcal{A}} = 0 \tag{6.31}$$

This choice is known as the Coulomb gauge. Using the Coulomb gauge Equation 6.30 reduces to

$$\nabla^2 \bar{\mathcal{A}} = -\mu_o \bar{\mathcal{J}} \tag{6.32}$$

which is known as the *vector Poisson's equation*. The general solution of this equation can be found by making use of the duality between the scalar Poisson's equation (Equation 6.7) and one vector component of the vector Poisson's equation:

$$\nabla^2 \mathcal{A}_i = -\mu_o J_i \tag{6.33}$$

where $i = x$, y, or z. We have already obtained the general solution to the scalar Poisson's equation (Equation 6.19), so it follows that a general solution to the vector Poisson's equation is

$$\bar{\mathcal{A}}(\bar{r}) = \frac{\mu_o}{4\pi} \int_{V'} \frac{\bar{\mathcal{J}}(\bar{r}')}{|\bar{r} - \bar{r}'|} dV' \tag{6.34}$$

6.2 Radiation from Time-Varying Charges

We now turn our attention to electromagnetic radiation from time-varying charges and currents. For ease of discussion, we repeat Maxwell's equations here for the time-varying case in a vacuum:

$$\nabla \times \bar{\mathcal{E}}(\bar{r}, t) = -\mu_o \frac{\partial \bar{\mathcal{H}}(\bar{r}, t)}{\partial t} \tag{6.35}$$

$$\nabla \times \bar{\mathcal{H}}(\bar{r}, t) = \varepsilon_o \frac{\partial \bar{\mathcal{E}}(\bar{r}, t)}{\partial t} + \bar{\mathcal{J}}(\bar{r}, t) \tag{6.36}$$

$$\nabla \cdot \bar{\mathcal{E}}(\bar{r}, t) = \frac{\rho_v(\bar{r}, t)}{\varepsilon_o} \tag{6.37}$$

$$\nabla \cdot \bar{\mathcal{H}}(\bar{r}, t) = 0 \tag{6.38}$$

6.2.1 The Inhomogeneous Wave Equations

For time-varying electric fields, the electric and magnetic fields are now coupled so that the curl of the electric field is no longer zero (Equation 6.35). Therefore the scalar potential, $\Phi(\bar{r})$, defined in Equation 6.6 is no longer valid. However, the divergence of the magnetic field is still equal to zero and therefore we can use the vector potential, $\bar{\mathcal{A}}$, defined in Equation 6.29, except that now $\bar{\mathcal{A}}$ is a time-varying quantity. Substituting $\bar{\mathcal{A}}$ into Faraday's law (Equation 6.35) yields:

$$\nabla \times \bar{\mathcal{E}}(\bar{r}, t) = -\mu_o \frac{\partial}{\partial t} \left\{ \frac{1}{\mu_o} \nabla \times \bar{\mathcal{A}}(\bar{r}, t) \right\} = -\nabla \times \frac{\partial \bar{\mathcal{A}}(\bar{r}, t)}{\partial t} \tag{6.39}$$

Collecting all of the terms we obtain

$$\nabla \times \left[\bar{\mathcal{E}}(\bar{r}, t) + \frac{\partial \bar{\mathcal{A}}(\bar{r}, t)}{\partial t} \right] = 0 \tag{6.40}$$

Since we again have obtained an expression in which the curl of a vector quantity is zero, we can define a new scalar potential

$$-\nabla\Phi(\overline{r},t) = \overline{\mathcal{E}}(\overline{r},t) + \frac{\partial\overline{\mathcal{A}}(\overline{r},t)}{\partial t} \tag{6.41}$$

Substituting $\Phi(\overline{r},t)$ and the vector potential $\overline{\mathcal{A}}(\overline{r},t)$ into Ampere's law (Equation 6.36), it can be shown that

$$\nabla^2\overline{\mathcal{A}} - \mu_o\varepsilon_o\frac{\partial^2\overline{\mathcal{A}}}{\partial t^2} = -\mu_o\overline{\mathcal{J}} + \nabla\left(\nabla\cdot\overline{\mathcal{A}} + \mu_o\varepsilon_o\frac{\partial\Phi}{\partial t}\right) \tag{6.42}$$

(refer to Appendix F for details). We are free to choose the divergence of $\overline{\mathcal{A}}$, so for convenience we choose

$$\nabla\cdot\overline{\mathcal{A}} = -\mu_o\varepsilon_o\frac{\partial\Phi}{\partial t} \tag{6.43}$$

which is known as the Lorentz gauge. Applying the Lorentz gauge, Equation 6.42 reduces to

$$\nabla^2\overline{\mathcal{A}} - \mu_o\varepsilon_o\frac{\partial^2\overline{\mathcal{A}}}{\partial t^2} = -\mu_o\overline{\mathcal{J}} \tag{6.44}$$

which is the *inhomogeneous vector wave equation.*

In a similar manner, we can substitute Equation 6.41 into Gauss' law (Equation 6.37) which yields (refer to Appendix F for details)

$$\nabla^2\Phi - \mu_o\varepsilon_o\frac{\partial^2\Phi}{\partial t^2} = -\frac{\rho_v}{\varepsilon_o} \tag{6.45}$$

which is the *inhomogeneous scalar wave equation.*

6.2.2 Solutions to the Inhomogeneous Wave Equations

If the sources $\rho_v(\overline{r},t)$ and $\overline{\mathcal{J}}(\overline{r},t)$ are time-harmonic, then we can make use of phasors. We make the following phasor transformations:

$$\begin{gathered}\Phi(\overline{r},t)\rightarrow\Phi(\overline{r})\\ \overline{\mathcal{A}}(\overline{r},t)\rightarrow\overline{A}(\overline{r})\\ \overline{\mathcal{J}}\left(\overline{r},t\right)\rightarrow\overline{J}(\overline{r})\\ \rho_v(\overline{r},t)\rightarrow\rho_v(\overline{r})\end{gathered}$$

which transform the inhomogeneous vector wave and scalar equations into the *inhomogeneous Helmholtz equations*

$$\nabla^2\overline{A}(\overline{r}) + \omega^2\mu_o\varepsilon_o\overline{A}(\overline{r}) = -\mu_o\overline{J}(\overline{r}) \tag{6.46}$$

$$\nabla^2\Phi(\overline{r}) + \omega^2\mu_o\varepsilon_o\Phi(\overline{r}) = -\frac{1}{\varepsilon_o}\rho_v(\overline{r}) \tag{6.47}$$

It can be shown that the potential created by an oscillating charge $q(t) = q_o\cos(\omega t)$ located at the origin (in free space) is

$$\Phi = \frac{q_o}{4\pi\varepsilon_o r}e^{-jk_o r} \tag{6.48}$$

(refer to Appendix F for details). Noting the similarities to the static case (see Equations 6.17 and 6.19), it follows that a time-harmonic continuous charge distribution at a location described by the position vector $\overline{r}'$ will give rise to the harmonic dynamic scalar potential

$$\Phi(\overline{r}) = \frac{1}{4\pi\varepsilon_o}\int_{V'}\frac{\rho_v(\overline{r}')}{|\overline{r}-\overline{r}'|}e^{-jk_o|\overline{r}-\overline{r}'|}dV' \tag{6.49}$$

Again noting the similarity between the scalar and vector equations (see Equation 6.34), the harmonic dynamic vector potential is

$$\bar{A}(\bar{r}) = \frac{\mu_o}{4\pi} \int_{V'} \frac{\bar{J}(\bar{r}')}{|\bar{r} - \bar{r}'|} e^{-jk_o|\bar{r}-\bar{r}'|} dV' \tag{6.50}$$

6.3 Radiation from a Hertzian Dipole

A Hertzian dipole consists of two interconnected charge reservoirs that contain equal and opposite amounts of charge. Assume that the charge reservoirs are positioned at $z = \pm d/2$ on the z-axis as shown in Figure 6.5, where d is much less than the electromagnetic wavelength. A general rule-of-thumb is that d should not exceed $\lambda_o/50$. Because $d \ll \lambda_o$, a Hertzian dipole is also called a "short" or an "infinitesimal" dipole. The charge oscillates between the charge reservoirs at frequency ω which gives rise to an oscillating current along the length of an "infinitely thin" conductor running between the charge reservoirs.

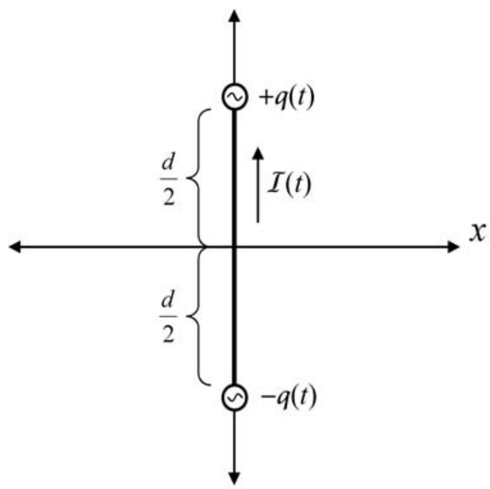


Figure 6.5: A Hertzian dipole, consisting of two interconnected charge reservoirs positioned a distance d apart on the z-axis. The charge oscillates between the reservoirs, giving rise to the oscillating current $\mathcal{I}(t)$.

Because the length of the infinitely thin conductor is much less than the wavelength, we can assume that at any instant in time the current is constant everywhere along the conductor. In this case, the time-harmonic current flowing in a Hertzian dipole can be written as

$$\mathcal{I}(t) = \frac{\partial}{\partial t} q(t) = I_o \cos(\omega t) \tag{6.51}$$

And the current density is

$$\bar{\mathcal{J}}(\bar{r}, t) = \frac{I_o \cos(\omega t)}{A_c} \hat{z}, \qquad -\frac{d}{2} \leq z \leq \frac{d}{2} \tag{6.52}$$

where A_c is the cross-sectional area of the conductor. In phasor form, the current density for the Hertzian dipole is then

$$\bar{J}(\bar{r}) = \frac{I_o}{A_c} \hat{z}, \qquad -\frac{d}{2} \leq z \leq \frac{d}{2} \tag{6.53}$$

In order to use Equation 6.50 to find the vector potential, we require the distance $|\bar{r} - \bar{r}'|$, which is the distance from our observation point to a small volume of integration along the conductor between the dipoles. However, as illustrated in Figure 6.6, since we are assuming that the dipole is short, we can approximate

$|\bar{r} - \bar{r}'| \approx |\bar{r}| = r$ everywhere along the dipole. In addition, the deferential volume element for the integration is $dV' = A_c dz$. Hence,

$$\bar{A}(\bar{r}) = \frac{\mu_o}{4\pi} \int_{-d/2}^{+d/2} \frac{(I_o/A_c)\,\hat{z}}{r} e^{-jk_o r} A_c dz = \frac{\mu_o I_o}{4\pi r} e^{-jk_o r} \int_{-d/2}^{+d/2} dz$$

$$= \hat{z} \frac{\mu_o I_o d}{4\pi r} e^{-jk_o r} \tag{6.54}$$

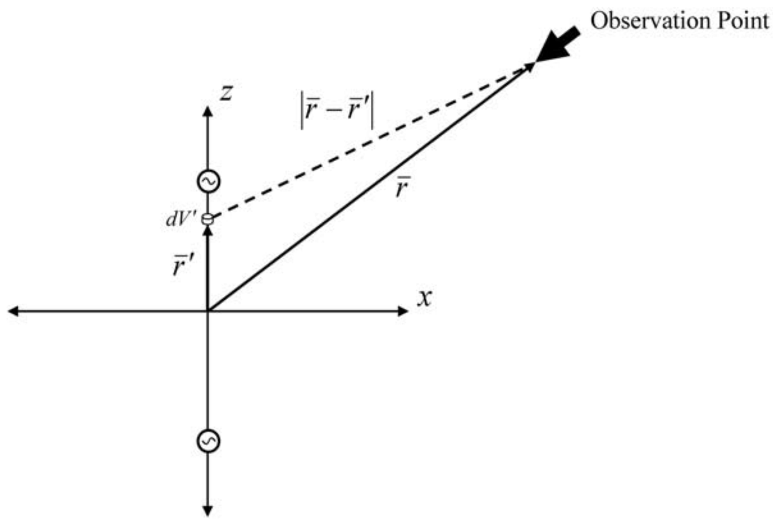


Figure 6.6: Since we are assuming that the Hertzian dipole is short, the distance $|\bar{r} - \bar{r}'|$ can be approximated by r, the distance from the observation point to the origin.

Because the dipole will radiate fields out from the origin, we convert to spherical coordinates. Using the relationship $\hat{z} = \hat{r}\cos\theta - \hat{\theta}\sin\theta$, the vector potential becomes

$$\bar{A}(r,\theta) = \left(\hat{r}\cos\theta - \hat{\theta}\sin\theta\right) \frac{\mu_o I_o d}{4\pi r} e^{-jk_o r} \tag{6.55}$$

The magnetic field intensity can now be calculated using Equation 6.29:

$$\bar{H}(r,\theta,\phi) = \frac{1}{\mu_o} \nabla \times \bar{A} = \frac{jk_o I_o d e^{-jk_o r}}{4\pi r} \left(1 + \frac{1}{jk_o r}\right) \sin\theta \hat{\phi} \tag{6.56}$$

(Refer to Chapter 1 for the curl of a vector field in spherical coordinates.) We then find $\bar{E}(r,\theta,\phi)$ from $\bar{H}(r,\theta,\phi)$ using Ampere's law (Equation 6.36) in phasor form, where $\bar{J}(\bar{r}) = 0$ for $r \neq 0$:

$$\bar{E}(r,\theta,\phi) = \frac{jk_o \eta_o I_o d}{4\pi r} e^{-jk_o r} \left\{ \left[\frac{1}{jk_o r} + \left(\frac{1}{jk_o r}\right)^2 \right] 2\cos\theta \hat{r} + \left[1 + \frac{1}{jk_o r} + \left(\frac{1}{jk_o r}\right)^2 \right] \sin\theta \hat{\theta} \right\} \tag{6.57}$$

The complex Poynting vector $\bar{S}_c = \bar{E} \times \bar{H}^*$ is then found from the electric and magnetic fields

$$\bar{S}_c = \eta_o \left(\frac{k_o I_o d}{4\pi r}\right)^2 \left\{ \left[1 + \left(\frac{1}{jk_o r}\right)^3 \right] \sin^2\theta \hat{r} - \left[\frac{1}{jk_o r} - \left(\frac{1}{jk_o r}\right)^3 \right] \sin 2\theta \hat{\theta} \right\} \tag{6.58}$$

The time-average power density, $\langle \bar{\mathcal{S}} \rangle$, is then

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{1}{2}\mathcal{R}e\left\{ \bar{S}_c \right\} = \frac{\eta_o}{2}\left(\frac{k_o I_o d}{4\pi r}\right)^2 \sin^2\theta \hat{r} \tag{6.59}$$

The total power radiated by the dipole is calculated by integrating the average power per unit area, $\langle \bar{\mathcal{S}} \rangle$, over the surface, S, of a sphere enclosing the dipole:

$$P = \oint_S \left(\left\langle \bar{\mathcal{S}} \right\rangle \cdot \hat{n} \right) dS = \int_0^{\pi} \int_0^{2\pi} \left[\frac{\eta_o}{2}\left(\frac{k_o I_o d}{4\pi r}\right)^2 \sin^2\theta \right] r^2 \sin\theta d\phi d\theta$$

$$= \frac{\eta_o}{12\pi}\left(k_o I_o d\right)^2 = \frac{\eta_o \mu_o \varepsilon_o}{12\pi}\left(\omega I_o d\right)^2 \tag{6.60}$$

Note that the total radiated power scales with the frequency squared.

Before we examine the fields of the Hertzian dipole in detail, let us examine the behavior of the $1/(k_o r)^n$ terms that are present in the field solutions. Figure 6.7 shows the behavior of $1/(k_o r)^n$ for $n = 1$, 2, and 3 as the distance from the dipole, r, is increased. From the plot, we see that close to the origin the $n = 3$ term dominates, but as the distance from the origin is increased, the $n = 1$ term dominates. The transition point occurs when $k_o r = 1$, or where $r = 1/k_o$. In order to simplify the expressions for the radiation, we will look at the fields in two limiting cases:

1. The region where $r \ll 1/k_o$, called the "near-field" zone. In this region, the higher powers of $1/(k_o r)^n$ dominate.
2. The region where $r \gg 1/k_o$, called the "far-field" zone. In this region, the lower powers of $1/(k_o r)^n$ dominate, and all of the $1/(k_o r)^n$ terms approach zero as $r \to \infty$.

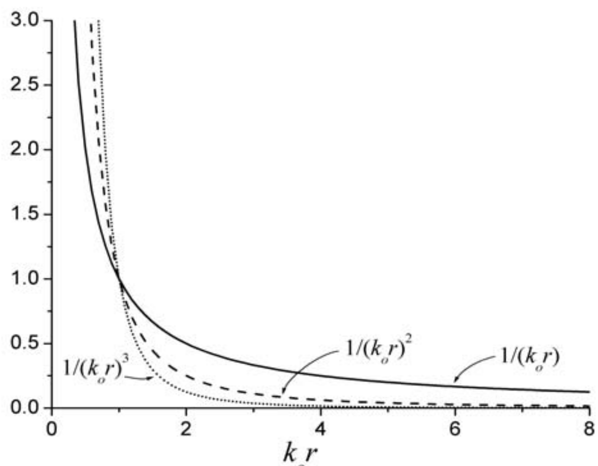


Figure 6.7: Behavior of the $1/(k_o r)^n$ terms of the Hertzian dipole radiation. Close to the origin the higher-power terms dominate, while far from the origin the lower-power terms dominate.

6.3.1 Near-field Characteristics

In the near field ($r \ll 1/k_o$), the highest powers of $1/(k_o r)^n$ are dominant. Therefore, we neglect all but the highest powers of $1/(k_o r)^n$ in the field expressions (Equations 6.56 and 6.57) to obtain approximate field expressions in the near-field:

$$\bar{H}_{\text{nf}} = \frac{I_o d}{4\pi r^2} e^{-jk_o r} \sin\theta\hat{\phi} \tag{6.61}$$

$$\bar{E}_{\text{nf}} = -\frac{jI_o d\eta_o e^{-jk_o r}}{4\pi r^3 k_o}\left(2\cos\theta\hat{r} + \sin\theta\hat{\theta}\right) \tag{6.62}$$

Now notice that we can convert the time-varying current and charge into phasor form as follows:

$$\mathcal{I}(t) = \frac{\partial q(t)}{\partial t} = I_o\cos(\omega t) \quad \Rightarrow \quad I = j\omega q = I_o \tag{6.63}$$

where I is the phasor form of $\mathcal{I}(t)$. From this we observe that $I_o = j\omega q$. Also, recall that the dipole moment is $p = qd$. We then use $k_o = \omega\sqrt{\mu_o\varepsilon_o}$ to write

$$I_o d = j\omega q d = j\omega p = \frac{jk_o p}{\sqrt{\mu_o\varepsilon_o}} \tag{6.64}$$

Finally, note that in the near-field $e^{-jk_o r} \to 1$. The electromagnetic fields in the near-field are then approximated as

$$\bar{H}_{\text{nf}} = \frac{j\omega p}{4\pi r^2}\sin\theta\hat{\phi} \tag{6.65}$$

$$\bar{E}_{\text{nf}} = \frac{p}{4\pi\varepsilon_o r^3}\left(2\cos\theta\hat{r} + \sin\theta\hat{\theta}\right) \tag{6.66}$$

Notice that, in the near-field solutions, the magnetic field is proportional to $1/r^2$, while the electric field is proportional to $1/r^3$. This means that the electric field will be dominant in the near-field (since r is small). Calculating the complex Poynting vector in the near-field we find

$$\bar{S}_{c,\text{nf}} = -\frac{jp^2\omega\left(\sin\theta\right)^2}{16\pi^2 r^5\varepsilon_o}\hat{r} + \frac{jp^2\omega\cos\theta\sin\theta}{8\pi^2 r^5\varepsilon_o}\hat{\theta} \tag{6.67}$$

The time-average power flow density in the near-field is then

$$\left\langle\bar{\mathcal{S}}_{\text{nf}}\right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{S}_{c,\text{nf}}\right\} = 0 \tag{6.68}$$

This indicates that there is no net power flow out of the near-field. Because the electric field dominates and there is no net power flow, electric energy is stored in the near-field. This means that the near-field behaves like a capacitor. Note that in the electrostatic limit ($\omega \to 0$), the magnetic field is zero and the electromagnetic field is identical to the electrostatic dipole. In this case, the fields behave like in a capacitor.

6.3.2 Far-field Characteristics

Typically, the radiators we use to generate electromagnetic waves are relatively distant from the location at which we are receiving the waves. For example, your "local" FM station may be ten miles away or more. For an FM radio station broadcasting at 100 MHz, $1/k_o \approx 0.5$ m, which means that we care much more about the far-field characteristics of the radiator.

In the far-field ($r \gg 1/k_o$), all of the $1/(k_o r)$ terms approach zero. Therefore, set $1/(k_o r) = 0$ in Equations 6.56 and 6.57 to obtain expressions for the magnetic and electric fields in the far-field:

$$\bar{H}_{\text{ff}} = \frac{jk_o I_o d}{4\pi r}e^{-jk_o r}\sin\theta\hat{\phi} \tag{6.69}$$

$$\bar{E}_{\text{ff}} = \eta_o\frac{jk_o I_o d}{4\pi r}e^{-jk_o r}\sin\theta\hat{\theta} \tag{6.70}$$

There are several important observations that can be made from this result. Most important is that at a fixed radius, the electric and magnetic fields are related by $\bar{H} = \hat{k} \times \bar{E}/\eta_o$, which is exactly the relationship between the fields in a plane wave! Therefore in the far-field the electromagnetic fields *behave like plane waves*. The near-field spatial dependence of $1/r^2$ and $1/r^3$ ensure that, for even moderate distances away from the radiator, the far-field approximations hold.

The complex Poynting vector in the far-field is

$$\bar{S}_{c,\text{ff}} = \bar{E} \times \bar{H}^* = \eta_o \left(\frac{k_o I_o d}{4\pi r}\right)^2 \sin^2\theta \hat{r} \tag{6.71}$$

The time-average power flow density in the far-field is then

$$\left\langle \bar{\mathcal{S}}_{\text{ff}} \right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{S}_{c,\text{ff}}\right\} = \frac{\eta_o}{2}\left(\frac{k_o I_o d}{4\pi r}\right)^2 \sin^2\theta\hat{r} = \frac{3P}{8\pi r^2}\sin^2\theta\hat{r} \tag{6.72}$$

where P is the total power radiated by the dipole (see Equation 6.60). Notice that this is the same expression for $\langle \bar{\mathcal{S}} \rangle$ as was calculated previously in Equation 6.59 using the exact field expressions.

6.4 Radiator Gain and Radiation Resistance

An electromagnetic radiator is often characterized by its gain, G, and its radiation resistance, R_{rad}. The *gain* of a radiator is defined as the ratio of the time-average power density (W/m^2) radiated in a particular direction (θ, ϕ) to the time-average power density radiated by an isotropic source emitting uniformly in all directions. For an isotropic radiating source emitting a total time-average power P, the time-average power density must be

$$\left\langle \bar{\mathcal{S}}_{\text{iso}} \right\rangle = \frac{P}{\text{spherical surface area}} = \frac{P}{4\pi r^2} \tag{6.73}$$

Therefore, for the Hertzian dipole, the gain is

$$G(\theta,\phi) = \frac{\left\langle \bar{\mathcal{S}}_{\text{dipole}} \right\rangle}{\left\langle \bar{\mathcal{S}}_{\text{iso}} \right\rangle} = \frac{3P\sin^2\theta/\left(8\pi r^2\right)}{P/\left(4\pi r^2\right)} = \frac{3}{2}\sin^2\theta \tag{6.74}$$

The *radiation pattern* of a radiator, $p(\theta, \phi)$, is defined as the gain $G(\theta, \phi)$ normalized such that the maximum value is unity. Therefore, the radiation pattern of the Hertzian dipole is simply

$$p(\theta,\phi) = \sin^2\theta \tag{6.75}$$

This radiation pattern is illustrated in Figure 6.8. In the figure, the dipole is oriented at the origin along the vertical axis. The radiation pattern for the dipole looks similar to a doughnut with a hole diameter that decreases to zero at the origin. Notice that there is *no radiation along the dipole axis*, and that the *maximum power is radiated perpendicular to the dipole axis.*

The *radiation resistance* of a radiator is defined as the equivalent resistor that, given the same time-average current, would dissipate the same power as is radiated by the radiator. Given a phasor current I flowing through a resistor, the time-average power dissipated in that resistor is $P = I^2_{\text{rms}}R = (|I|/\sqrt{2})^2 R$. Solving for R, we find that the equivalent radiation resistance of a radiator is

$$R_{\text{rad}} = \frac{P}{|I|^2/2} \tag{6.76}$$

The radiation resistance is a measure of how effectively the antenna converts a current signal into radiated power. Note that the radiation resistance is *not* a measure of loss in the antenna; for example, the ideal

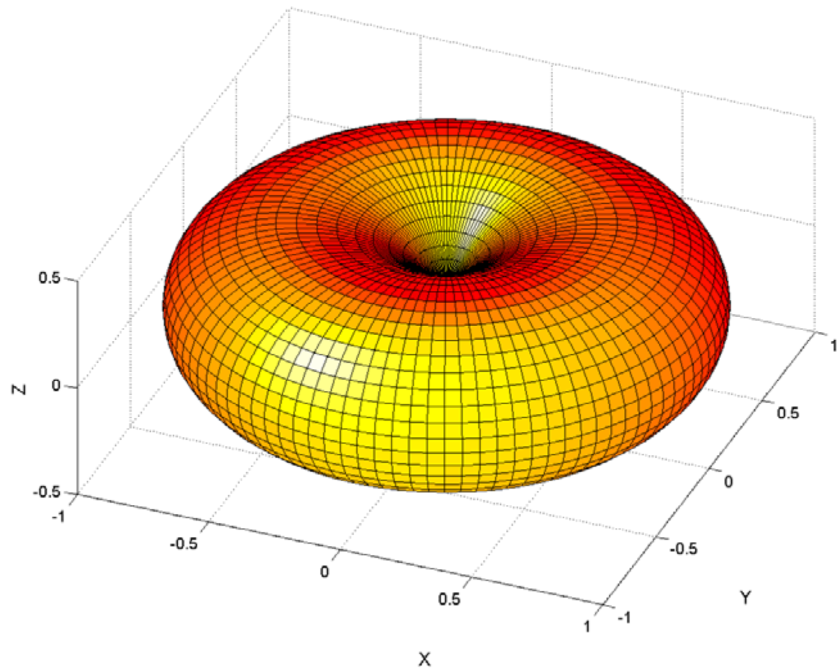


Figure 6.8: Radiation pattern of a Hertzian dipole. The dipole is oriented at the origin along the z-axis, which is the vertical axis in the figure.

Hertzian dipole is lossless but has a nonzero radiation resistance. For a Hertzian dipole, the radiation resistance is

$$R_{\text{rad}} = \frac{\eta_o (k_o I_o d)^2 / 12\pi}{I_o^2/2} = \frac{(dk_o)^2\,\eta_o}{6\pi} \approx 20\,(dk_o)^2 \approx 790\left(\frac{d}{\lambda_o}\right)^2 \tag{6.77}$$

In a Hertzian dipole, the distance d is much smaller than the wavelength. A typical value is $d/\lambda_o = 0.01$, which results in a radiation resistance of $R_{\text{rad}} \approx 0.08\Omega$. Keeping in mind that free space has an impedance of $\eta_o \approx 377\Omega$, a Hertzian dipole is a poor radiator.

6.5 Radiation from Arrays of Hertzian Dipoles

As stated previously, when dealing with radiators we are most often interested in the far-field radiation characteristics. Since superposition holds, the electric fields of individual radiators add in the far-field. Since the electric field has both a magnitude and a phase, the radiation pattern can be modified by controlling not only the geometric arrangement of the various radiators, but also the relative phase between them.

Let us examine the far-field electric field of Hertzian dipole in more detail so that we can gain an appreciation for how the fields will interfere constructively or destructively when multiple dipoles are present. We express the far-field electric field as

$$\bar{E}_{\text{ff}} = \eta_o \frac{jk_o I_o d}{4\pi r} e^{-jk_o r} \sin\theta\hat{\theta} = \frac{E_o}{r} e^{-jk_o r} \sin\theta\hat{\theta} \tag{6.78}$$

where we have defined the arbitrary constant $E_o = j\eta_o k_o I_o d/(4\pi)$. The time-domain electric field is then

$$\bar{\mathcal{E}}_{\text{ff}} = \mathcal{R}e\left\{\bar{E}_{\text{ff}} e^{j\omega t}\right\} = \frac{E_o}{r}\cos(\omega t - k_o r)\sin\theta\hat{\theta} \tag{6.79}$$

Notice the $1/r$ dependence in the far-field approximation. Obviously, this approximation does not hold close to the dipole or else $\bar{\mathcal{E}}_{\rm ff}$ would approach infinity as $r \to 0$! A plot of $\bar{\mathcal{E}}_{\rm ff}$ at $t = 0$ and $\theta = 90°$ (the x-y plane, see Figure 6.1) is shown below in Figure 6.9. This field pattern will propagate radially as time progresses. Note the positive and negative field values.

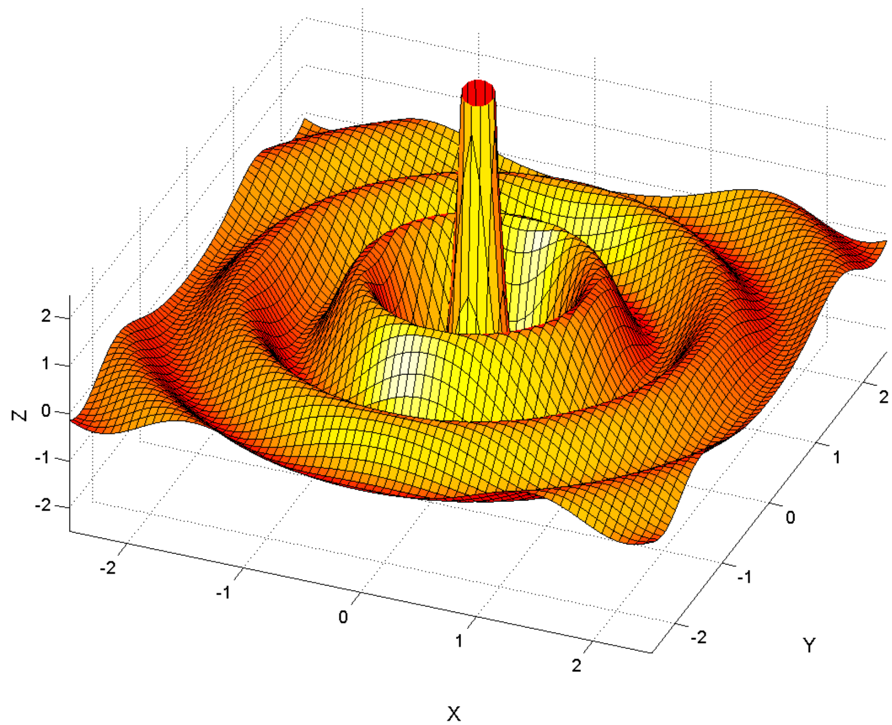


Figure 6.9: Time-domain far-field electric field pattern from a Hertzian dipole at $t = 0$ and $\theta = 90°$ (x-y plane, which is perpendicular to the dipole axis).

As an example of a simple antenna array consider two Hertzian dipoles positioned at $x = +D/2$ and $x = -D/2$, with the axis of both dipoles aligned with the z-axis, as shown in Figure 6.10. Realistically, this would correspond to a situation in which we have two dipole antennas oriented vertically with respect to the ground (x-y plane) spaced a distance D apart. Dipole 1 has a length d_1 and carries a phasor current I_1. Dipole 2 has a length d_2 and carries a phasor current I_2. The currents I_1 and I_2 most likely have different amplitudes and are out-of-phase. We want to find the total electric field in the far-field for these two dipoles.

Figure 6.11 illustrates the two-dipole array in the x-y plane (the view from "above" the dipoles). The distance r_1 is the distance from the observation point to dipole 1, and the distance r_2 is the distance to dipole 2. By superposition, we then find the total electric field of the two Hertzian dipoles in the far-field by summing the far-field electric fields of two single Hertzian dipoles:

$$\bar{E}_{\rm ff} = \eta_o \frac{jk_o}{4\pi}\left(\frac{d_1 I_1}{r_1}\sin\theta_1 e^{-jk_o r_1} + \frac{d_2 I_2}{r_2}\sin\theta_2 e^{-jk_o r_2}\right)\hat{\theta} \tag{6.80}$$

Since we are in the far field, we assume that the spacing between the dipoles is much less than the distances r_1 and r_2. Therefore $1/r_1 \approx 1/r_2 \approx 1/r$, where r is the distance to the origin. Using the same argument, $\theta_1 \approx \theta_2 \approx \theta$. We cannot simply approximate $r_1 \approx r_2$ in the exponential terms, however, because these phase

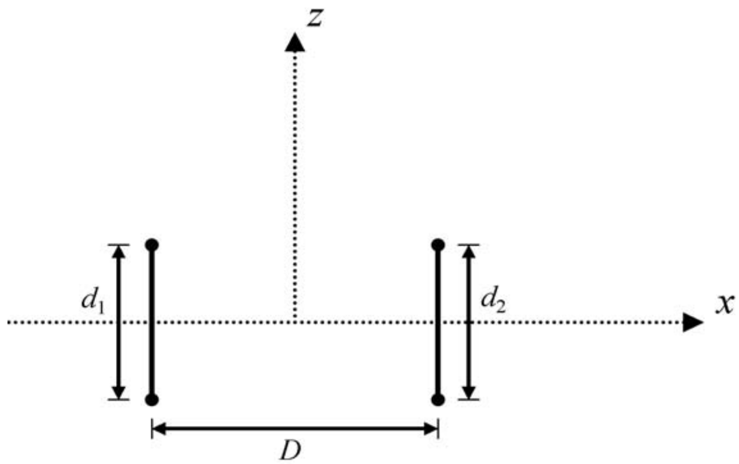


Figure 6.10: Two Hertzian dipoles, spaced a distance D apart on the x-axis.

terms vary rapidly even with very small changes in r. In this case, we use a similar approximation technique as we used previously for the separated static charges (see Figure 6.11):

$$r_1 \approx r + \frac{D}{2}\cos\phi \tag{6.81}$$

and

$$r_2 \approx r - \frac{D}{2}\cos\phi \tag{6.82}$$

Using these approximations, the total electric field is written as

$$\bar{E}_{\text{ff}} = \eta_o \frac{jk_o d_1 I_1}{4\pi r}\sin\theta e^{-jk_o(r+\frac{D}{2}\cos\phi)}\left(1+\frac{d_2 I_2}{d_1 I_1}e^{jk_o D\cos\phi}\right)\hat{\theta} \tag{6.83}$$

We now define

$$Ae^{j\alpha} = \frac{d_2 I_2}{d_1 I_1} \tag{6.84}$$

which yields

$$\bar{E}_{\text{ff}} = \eta_o \frac{jk_o d_1 I_1}{4\pi r}\sin\theta e^{-jk_o(r+\frac{D}{2}\cos\phi)}\left(1+Ae^{j\alpha}e^{jk_o D\cos\phi}\right)\hat{\theta} \tag{6.85}$$

Since the far-field radiation behaves like plane waves radiating from the origin, the time-average power $\langle\bar{\mathcal{S}}\rangle$ is proportional to $\bar{E}_{\text{ff}}\cdot\bar{E}^*_{\text{ff}}$:

$$\begin{aligned}\left\langle\bar{\mathcal{S}}\right\rangle &= \frac{S_o}{r^2}\sin^2\theta\left(1+Ae^{j\alpha}e^{jk_o D\cos\phi}\right)\left(1+Ae^{-j\alpha}e^{-jk_o D\cos\phi}\right)\\ &= \frac{S_o}{r^2}\sin^2\theta[1+A^2+2A\cos(\alpha+k_o D\cos\phi)]\end{aligned} \tag{6.86}$$

where S_o includes all of the constant terms. The gain of the dipole array is

$$G(\theta,\phi) = \frac{\left\langle\bar{\mathcal{S}}_{\text{dipoles}}\right\rangle}{\left\langle\bar{\mathcal{S}}_{\text{iso}}\right\rangle} = \frac{\left(S_o/r^2\right)\sin^2\theta[1+A^2+2A\cos(\alpha+k_o D\cos\phi)]}{P/\left(4\pi r^2\right)}$$

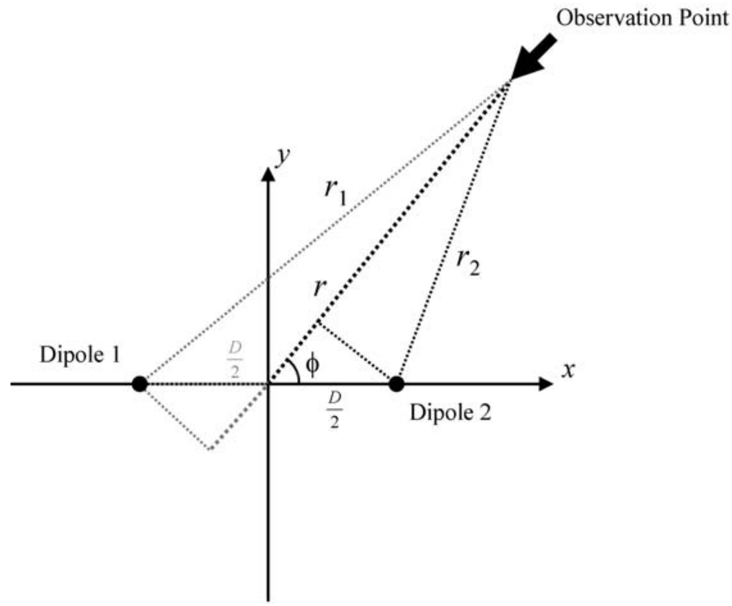


Figure 6.11: Illustration of the two-dipole array in the x-y plane.

$$= K \sin^2 \theta [1 + A^2 + 2A \cos(\alpha + Dk_o \cos\phi)] \tag{6.87}$$

where K includes all of the constant terms. The exact value of K is not important since, to obtain the radiation pattern, we normalize $G(\theta, \phi)$ to have a maximum value of unity. The normalized radiation pattern is then

$$p(\theta, \phi) = \frac{[1 + A^2 + 2A \cos(\alpha + Dk_o \cos\phi)] \sin^2 \theta}{(1 + A)^2} \tag{6.88}$$

Using this general expression for the radiation pattern, we will look at five specific cases. Remember that A is the magnitude of the ratio $d_2 I_2 / d_1 I_1$, α is the phase difference between the two dipoles, and D is the spacing between the two dipoles.

Case 1: $A = 1$, $\alpha = 0$, and $D = \lambda_o/2$ (equal amplitude dipoles, spaced $\lambda_o/2$ apart, and radiating in-phase)

The radiation pattern is

$$p(\theta, \phi) = \left[1 + 1 + 2\cos\left(\frac{\lambda_o}{2} \frac{2\pi}{\lambda_o} \cos\phi\right)\right] \sin^2\theta / (1 + 1)^2$$

$$= \frac{[1 + \cos(\pi \cos\phi)] \sin^2 \theta}{2} \tag{6.89}$$

The pattern is shown below in Figure 6.12. Notice the two lobes oriented perpendicular to the x-axis, i.e. perpendicular to the axis connecting the two dipoles.

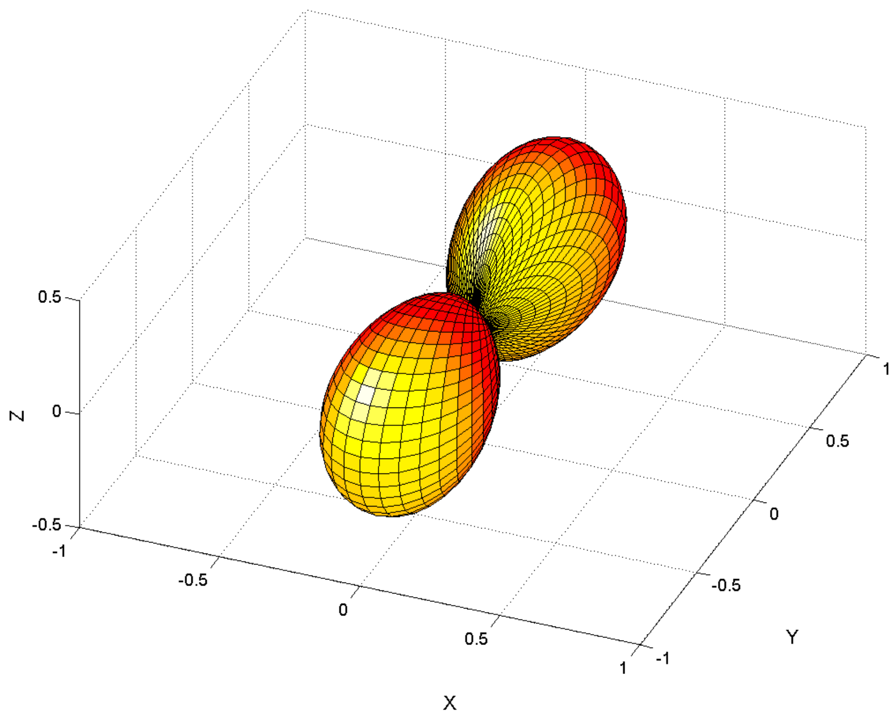


Figure 6.12: Radiation pattern with $A = 1$, $\alpha = 0$, and $D = \lambda_o/2$.

Case 2: $A = 1$**,** $\alpha = \pi$**, and** $D = \lambda_o/2$ **(equal amplitude dipoles, spaced $\lambda_o/2$ apart, and radiating 180 degrees out-of-phase)**

The radiation pattern is

$$p(\theta, \phi) = \left[1 + 1 + 2\cos\left(\pi + \frac{\lambda_o}{2}\frac{2\pi}{\lambda_o}\cos\phi\right)\right]\sin^2\theta/\left(1+1\right)^2$$
$$= \frac{[1 - \cos(\pi\cos\phi)]\sin^2\theta}{2} \tag{6.90}$$

This is shown in Figure 6.13. Notice that now the lobes are oriented along the x-axis (parallel to the axis connecting the two dipoles).

Case 3: $A = 1$**,** $\alpha = \pi$**, and** $D = 3\lambda_o/2$ **(equal amplitude dipoles, spaced $3\lambda_o/2$ apart, and radiating 180 degrees out-of-phase)**

The radiation pattern is

$$p(\theta, \phi) = \left[1 + 1 + 2\cos\left(\pi + \frac{3\lambda_o}{2}\frac{2\pi}{\lambda_o}\cos\phi\right)\right]\sin^2\theta/\left(1+1\right)^2$$
$$= \frac{[1 - \cos(3\pi\cos\phi)]\sin^2\theta}{2} \tag{6.91}$$

The radiation pattern now shows six distinct lobes (Figure 6.14).

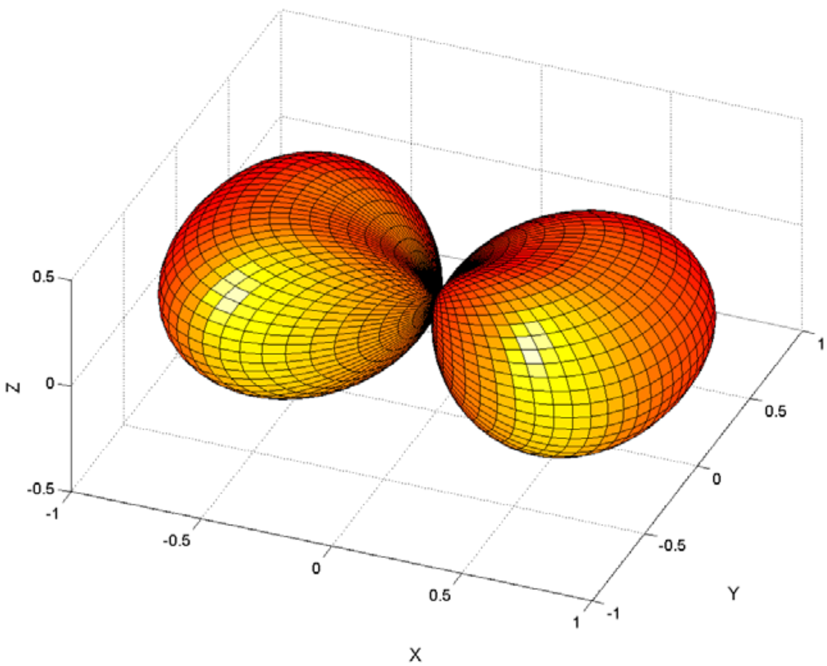


Figure 6.13: Radiation pattern with $A = 1$, $\alpha = \pi$, and $D = \lambda_o/2$.

Case 4: $A = 1$, $\alpha = \pi/2$, **and** $D = \lambda_o/2$ **(equal amplitude dipoles, spaced $\lambda_o/2$ apart, and radiating 90 degrees out-of-phase)**

The radiation pattern is

$$p(\theta, \phi) = \left[1 + 1 + 2\cos\left(\frac{\pi}{2} + \frac{\lambda_o}{2}\frac{2\pi}{\lambda_o}\cos\phi\right)\right]\sin^2\theta / (1+1)^2$$

$$= \frac{[1 - \sin(\pi\cos\phi)]\sin^2\theta}{2} \tag{6.92}$$

The pattern has two main lobes and one smaller lobe (Figure 6.15).

Case 5: $A = 2$, $\alpha = 0$, **and** $D = \lambda_o$ **(unequal amplitude dipoles, spaced λ_o apart, and radiating in-phase)**

The radiation pattern is

$$p(\theta, \phi) = \left[1 + 4 + 4\cos\left(\lambda_o\frac{2\pi}{\lambda_o}\cos\phi\right)\right]\sin^2\theta / (1+2)^2$$

$$= \frac{[5 + 4\cos(2\pi\cos\phi)]\sin^2\theta}{9} \tag{6.93}$$

The pattern is shown in Figure 6.16.

≪≪≪≪≪≪≪≫≫≫≫≫≫≫

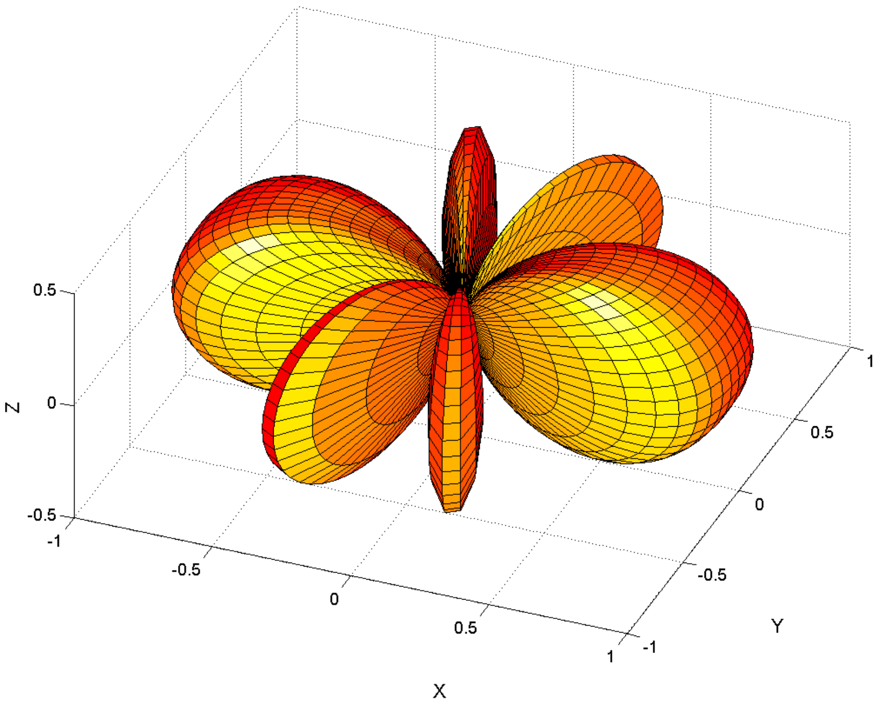


Figure 6.14: Radiation pattern with $A = 1$, $\alpha = \pi$, and $D = 3\lambda_o/2$.

EXAMPLE 6.1:

Assume that you are using a two-dipole antenna array to transmit a 1 MHz AM signal. Both antennas are 6 m tall. Dipole 2 is east of dipole 1 by a distance D. The current flowing in dipole 1 is $\mathcal{I}_1(t) = 5\cos(\omega t)$ A, and the current in dipole 2 is $\mathcal{I}_2(t) = 5\cos(\omega t + 90°)$ A. What is the minimum spacing D (in meters) so that no signal is broadcast to the east?

Solution:

We first set up a coordinate system that corresponds to the directions north, south, east, and west. We are concerned with radiation along the earth's surface, so we focus on the x-y plane ($\theta = 90°$, see Figure 6.1). For convenience, we choose east to be the +x-direction ($\phi = 0°$) and north to be the +y-direction ($\phi = 90°$), as shown in Figure 6.17. The phasor currents for the two dipoles are $I_1 = 5$ A and $I_2 = 5e^{j\pi/2}$ A, which yields

$$Ae^{j\alpha} = \frac{d_2 I_2}{d_1 I_1} = \frac{(6)\left(5e^{j\pi/2}\right)}{(6)\,(5)} = e^{j\pi/2}$$

From this we see that $A = 1$ and $\alpha = \pi/2$. Notice that the individual lengths of the dipoles do not matter, only the ratio. The same is true of the current amplitudes. We are interested in the antenna pattern to the east, so we substitute $\theta = 90°$ and $\phi = 0°$ into the antenna pattern (Equation 6.88), which then reduces to

$$p(90°, 0°) = \frac{\left\{1 + 1^2 + 2\,(1)\cos[\pi/2 + Dk_o\cos(0°)]\right\}\sin^2(90°)}{(1+1)^2}$$

$$= \frac{1}{2} + \frac{1}{2}\cos(\pi/2 + Dk_o)$$

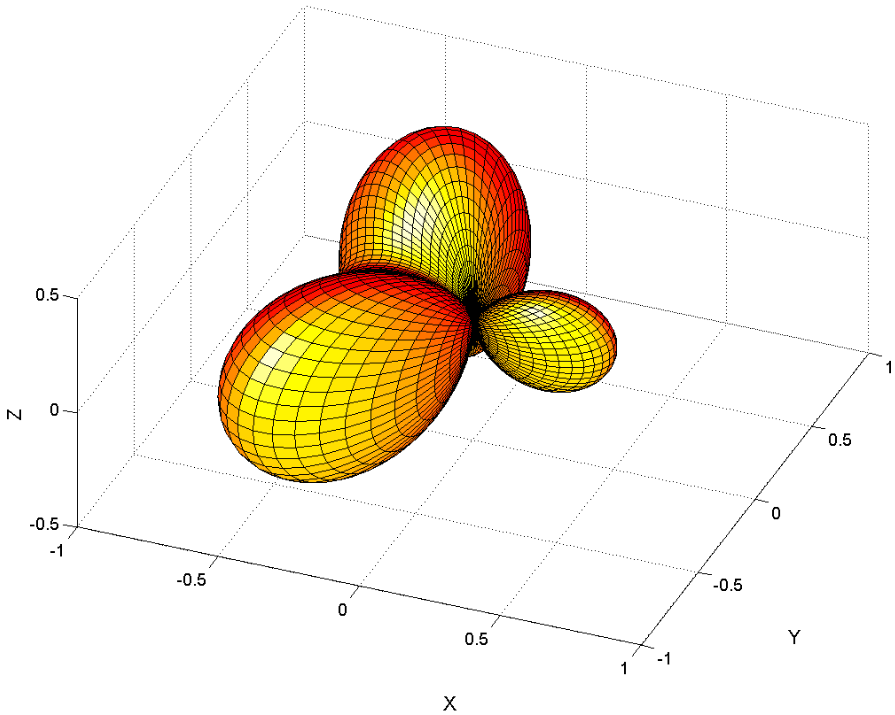


Figure 6.15: Radiation pattern with $A = 1$, $\alpha = \pi/2$, and $D = \lambda_o/2$.

where $k = 2\pi/\lambda$. This will equal zero only if $\cos(\pi/2 + Dk_o) = -1$. The solution for D then proceeds as follows:

$$\cos(\pi/2 + 2\pi D/\lambda_o) = -1$$

$$\pi/2 + 2\pi D/\lambda_o = ... - 3\pi, -\pi, \pi, 3\pi...$$

$$1/2 + 2D/\lambda_o = ... - 3, -1, 1, 3...$$

$$2D/\lambda_o = ... - \frac{7}{2}, -\frac{3}{2}, \frac{1}{2}, \frac{5}{2}...$$

$$D/\lambda_o = ... - \frac{7}{4}, -\frac{3}{4}, \frac{1}{4}, \frac{5}{4}...$$

Which value do we choose? Since D cannot be negative and we want the smallest possible distance, we choose $D/\lambda_o = 1/4$. At a frequency of 1 MHz, the wavelength in free space is $\lambda_o = c/f = 3 \times 10^8/10^6 = 300$ m. Therefore the distance D in meters is

$$D = \frac{\lambda_o}{4} = \frac{300}{4} = 75\,\text{meters}$$

The radiation pattern for $A = 1$, $\alpha = \pi/2$, and $D = \lambda_o/4$ is shown in Figure 6.18. Notice that all of the radiation is confined in the -x-direction, which is to the west in our example. This type of pattern is called a "cardioid."

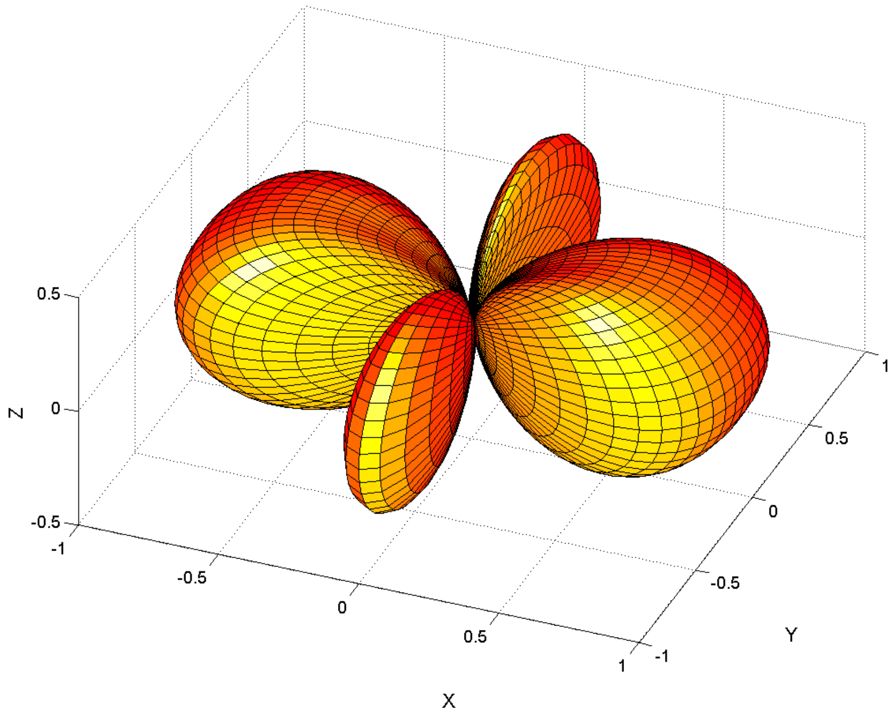


Figure 6.16: Radiation pattern with $A = 2$, $\alpha = 0$, and $D = \lambda_o$.

6.6 Important Concepts

Upon the completion of this chapter, you should be able to:

- Describe the difference between the scalar and vector potential, and why these potentials are introduced in the electromagnetic field solution.
- Understand the derivation of the inhomogeneous wave equations and the solutions to these equations.
- For a single Hertzian Dipole:
  - Derive the electromagnetic fields that are radiated by the dipole.
  - Understand the difference between the "near-field" and "far-field" regions.
  - Describe the fields and power flow in the near- and far-field regions.
- Understand the concepts of radiator gain and radiation resistance.
- Calculate and interpret radiation patterns.
- Understand the calculation of radiation patterns for arrays of two Hertzian dipoles.

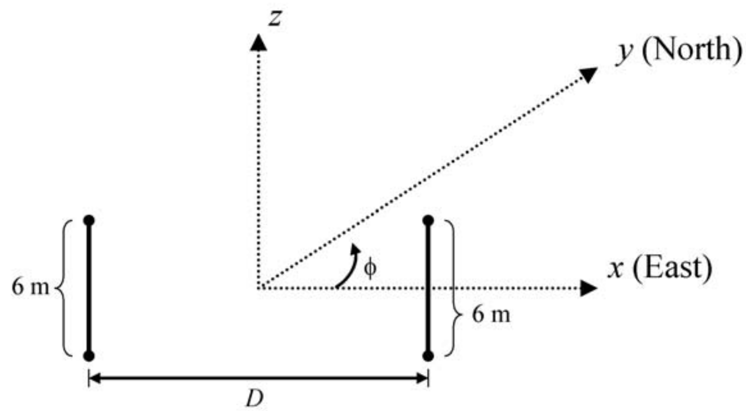


Figure 6.17: Dipole array for Example 6.1.

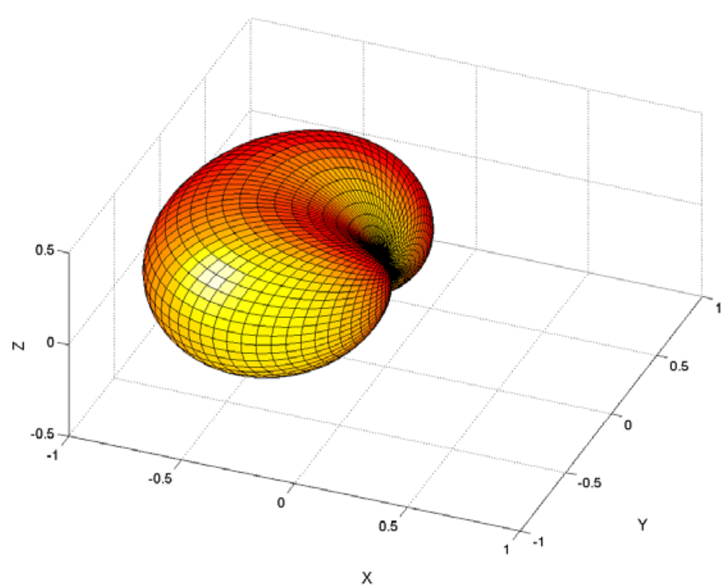


Figure 6.18: Radiation pattern for antenna array of Example 6.1 ($A = 1$, $\alpha = \pi/2$, and $D = \lambda_o/4$).

Appendix A

Derivations for Chapter 1

A.1 The Average of the Cross Product of Two Vectors

Following is the proof that the average of the cross product of two vectors can be directly calculated from their phasor representations. First, separate $\bar{\mathcal{A}}(t) = \mathcal{R}e\{\bar{A}e^{j\omega t}\}$ into real and imaginary parts

$$\bar{\mathcal{A}}(t) = \mathcal{R}e\left\{\left(\bar{A}_R + j\bar{A}_I\right)[\cos(\omega t) + j\sin(\omega t)]\right\}$$

$$= \bar{A}_R\cos(\omega t) - \bar{A}_I\sin(\omega t) \tag{A.1}$$

where $\bar{A}_R$ and $\bar{A}_I$ are the real and imaginary portions of the vector phasor $\bar{A}$, respectively. Likewise, $\bar{\mathcal{B}}(t) = \bar{B}_R\cos(\omega t) - \bar{B}_I\sin(\omega t)$. The cross product is

$$\bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t) = \left[\bar{A}_R\cos(\omega t) - \bar{A}_I\sin(\omega t)\right] \times \left[\bar{B}_R\cos(\omega t) - \bar{B}_I\sin(\omega t)\right]$$

$$= \left(\bar{A}_R \times \bar{B}_R\right)\cos^2(\omega t) + \left(\bar{A}_I \times \bar{B}_I\right)\sin^2(\omega t)$$

$$-\frac{1}{2}\left(\bar{A}_R \times \bar{B}_I + \bar{A}_I \times \bar{B}_R\right)\sin(2\omega t) \tag{A.2}$$

We have already shown that $\langle\sin(2\omega t)\rangle = 0$, so the average value of the cross product is simply

$$\left\langle\bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t)\right\rangle = \left\langle\left(\bar{A}_R \times \bar{B}_R\right)\cos^2(\omega t) + \left(\bar{A}_I \times \bar{B}_I\right)\sin^2(\omega t)\right\rangle \tag{A.3}$$

This relationship is further simplified by noting that $\langle\cos(\omega t)^2\rangle = \langle\sin(\omega t)^2\rangle = 1/2$, and that $\bar{A}_R$, $\bar{A}_I$, $\bar{B}_R$, and $\bar{B}_I$ are constant vectors, so the time average reduces to

$$\left\langle\bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t)\right\rangle = \frac{1}{2}\left(\bar{A}_R \times \bar{B}_R + \bar{A}_I \times \bar{B}_I\right) \tag{A.4}$$

Now note the following:

$$\frac{1}{2}\left(\bar{A} \times \bar{B}^*\right) = \frac{1}{2}\left(\bar{A}_R + j\bar{A}_I\right) \times \frac{1}{2}\left(\bar{B}_R - j\bar{B}_I\right)$$

$$= \frac{1}{2}\left(\bar{A}_R \times \bar{B}_R + \bar{A}_I \times \bar{B}_I\right) + j\frac{1}{2}\left(\bar{A}_I \times \bar{B}_R - \bar{A}_R \times \bar{B}_I\right) \tag{A.5}$$

So the real part of $\frac{1}{2}(\bar{A} \times \bar{B}^*)$ equals $\langle\bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t)\rangle$, and therefore

$$\left\langle\bar{\mathcal{A}}(t) \times \bar{\mathcal{B}}(t)\right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{A} \times \bar{B}^*\right\}. \tag{A.6}$$

Appendix B

Derivations for Chapter 2

B.1 Time-Harmonic Form of Maxwell's Equations

To convert Maxwell's equations into time-harmonic form, note that the time-domain fields are related to the phasor fields as follows:

$$\bar{\mathcal{E}}\left(\bar{r},t\right) = \mathcal{R}e\left\{\bar{E}e^{j\omega t}\right\}, \quad \bar{\mathcal{B}}(\bar{r},t) = \mathcal{R}e\left\{\bar{B}e^{j\omega t}\right\} \tag{B.1}$$

Substitute these into the curl equation $\nabla \times \bar{\mathcal{E}}(\bar{r},t) = -\partial\bar{\mathcal{B}}(\bar{r},t)/\partial t$:

$$\nabla \times \bar{\mathcal{E}}\left(\bar{r},t\right) = \mathcal{R}e\left\{\nabla \times \bar{E}e^{j\omega t}\right\} = \mathcal{R}e\left\{-j\omega\bar{B}e^{j\omega t}\right\} \tag{B.2}$$

so therefore

$$\mathcal{R}e\left\{\left(\nabla \times \bar{E} + j\omega\bar{B}\right)e^{j\omega t}\right\} = 0 \tag{B.3}$$

This equation must hold for all time, so we have $\mathcal{R}e\{\nabla \times \bar{E} + j\omega\bar{B}\} = 0$. At time $t = t_o + \pi/2\omega$ the equation yields

$$\begin{aligned} &\mathcal{R}e\left\{\nabla \times \bar{E}e^{j\pi/2}e^{j\omega t_o} + j\omega\bar{B}e^{j\pi/2}e^{j\omega t_o}\right\} \\ &= \mathcal{R}e\left\{j\left(\nabla \times \bar{E} + j\omega\bar{B}\right)e^{j\omega t_o}\right\} = 0 \end{aligned} \tag{B.4}$$

This equation also must hold for all time (any value of t_o). In addition, for a complex number $c = a + jb$, we have $\mathcal{R}e\{jc\} = -b = -\mathcal{I}m\{c\}$, so it follows that

$$\mathcal{I}m\left\{\nabla \times \bar{E} + j\omega\bar{B}\right\} = 0 \tag{B.5}$$

Combining the restriction on the real part and the imaginary part then it follows that

$$\nabla \times \bar{E} + j\omega\bar{B} = 0 \tag{B.6}$$

Analogous arguments are used to convert the remaining equations.

B.2 General Solution of the Helmholtz Equation

Substituting the phasor field $\bar{E} = E_x\hat{x} + E_y\hat{y} + E_z\hat{z}$ into the Helmholtz equation and separating the x-, y- and z-components, we obtain

$$\begin{aligned} \nabla^2 E_x + k^2 E_x &= 0 \\ \nabla^2 E_y + k^2 E_y &= 0 \\ \nabla^2 E_z + k^2 E_z &= 0 \end{aligned} \tag{B.7}$$

To solve these partial differential equations, apply the method of separation of variables. Using the x-component as an example, assume that the electric field component can be written as a product of three functions $f(x)$, $g(y)$ and $h(z)$, i.e.

$$E_x = f(x)g(y)h(z) \tag{B.8}$$

Substituting this field in the Helmholtz equation for the x-component gives $(\partial^2 f/\partial x^2)gh + f(\partial^2 g/\partial y^2)h + fg(\partial^2 h/\partial z^2) + k^2 fgh = 0$ which reduces to

$$\frac{\partial^2 f}{\partial x^2}\frac{1}{f(x)} + \frac{\partial^2 g}{\partial y^2}\frac{1}{g(y)} + \frac{\partial^2 h}{\partial z^2}\frac{1}{h(z)} + k^2 = 0 \tag{B.9}$$

Note that the first term will be a function of x, the second a function of y, the third term a function of z, and the fourth term is a constant. The only way that this relationship can be satisfied is if each term is equal to a constant, such that

$$\frac{\partial^2 f}{\partial x^2}\frac{1}{f(x)} = -k_x^2, \quad \frac{\partial^2 g}{\partial y^2}\frac{1}{g(y)} = -k_y^2, \quad \frac{\partial^2 h}{\partial z^2}\frac{1}{h(z)} = -k_z^2 \tag{B.10}$$

and $k_x^2 + k_y^2 + k_z^2 = k^2$. The solutions to these equations are as follows:

$$\begin{aligned} f(x) &= a_1 e^{-jk_x x} + a_2 e^{jk_x x} \\ g(y) &= b_1 e^{-jk_y y} + b_2 e^{jk_y y} \\ h(z) &= c_1 e^{-jk_z z} + c_2 e^{jk_z z} \end{aligned} \tag{B.11}$$

Remember that $E_x = f(x)g(y)h(z)$, so therefore

$$E_x = A_1 e^{-j\bar{k}\cdot\bar{r}}\hat{x} + A_2 e^{+j\bar{k}\cdot\bar{r}}\hat{x} \tag{B.12}$$

The first term is a wave propagating in the $+\bar{k}$ direction, and the second is propagating in the opposite direction. Similar solutions can be obtained for the y- and z-components, which can be combined to yield

$$\bar{E} = \bar{A}_1 e^{-j\bar{k}\cdot\bar{r}} + \bar{A}_2 e^{+j\bar{k}\cdot\bar{r}} \tag{B.13}$$

Another way to find the solution is to show that $\bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$ is a solution of the Helmholtz equation. Since solutions to partial differential equations are unique, this then proves the solution. First, divide the Helmholtz vector equation into 3 scalar equations: one for E_x, one for E_y, and another for E_z. For example, applying the x-component $E_x = E_{\text{xo}} e^{-j(k_x x + k_y y + k_z z)}$ and evaluating

$$\nabla^2 E_x = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) E_{\text{xo}} e^{-j(k_x x + k_y y + k_z z)} \tag{B.14}$$

yields

$$\nabla^2 E_x = -\left(k_x^2 + k_y^2 + k_z^2\right) E_{\text{xo}} e^{-j(k_x x + k_y y + k_z z)} = -k^2 E_x \tag{B.15}$$

which means

$$\nabla^2 E_x + k^2 E_x = 0 \tag{B.16}$$

Similar proofs are used for the remaining E_y and E_z components.

B.3 Orthogonality of Wave Vector and Electric Field

First, rewrite $\nabla \cdot \bar{E}$ as $\nabla \cdot \bar{E}_o e^{-j\bar{k}\cdot\bar{r}}$ where $\bar{E}_o = E_x\hat{x} + E_y\hat{y} + E_z\hat{z}$ is the amplitude of the wave. Then, we have

$$\nabla \cdot \bar{E} = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \right) \cdot \left(E_x e^{-j\bar{k}\cdot\bar{r}}\hat{x} + E_y e^{-j\bar{k}\cdot\bar{r}}\hat{y} + E_z e^{-j\bar{k}\cdot\bar{r}}\hat{z} \right) \tag{B.17}$$

and recall that $\bar{k} \cdot \bar{r} = k_x x + k_y y + k_z y$. Note that now E_x, E_y, and E_z are no longer functions of position. Evaluating the dot product then yields:

$$\begin{aligned} \nabla \cdot \bar{E} &= \frac{\partial}{\partial x} E_x e^{-j\bar{k}\cdot\bar{r}}\hat{x} + \frac{\partial}{\partial y} E_y e^{-j\bar{k}\cdot\bar{r}}\hat{y} + \frac{\partial}{\partial z} E_z e^{-j\bar{k}\cdot\bar{r}}\hat{z} \\ &= -jk_x E_x e^{-j\bar{k}\cdot\bar{r}} - jk_y E_y e^{-j\bar{k}\cdot\bar{r}} - jk_z E_z e^{-j\bar{k}\cdot\bar{r}} \\ &= -j\left(k_x\hat{x} + k_y\hat{y} + k_z\hat{z}\right) \cdot \left(E_x e^{-j\bar{k}\cdot\bar{r}}\hat{x} + E_y e^{-j\bar{k}\cdot\bar{r}}\hat{y} + E_z e^{-j\bar{k}\cdot\bar{r}}\hat{z} \right) \\ &= -j\bar{k} \cdot \bar{E} \end{aligned} \tag{B.18}$$

and since in free space $\nabla \cdot \bar{E} = 0$ it follows that

$$\bar{k} \cdot \bar{E} = 0$$

Therefore $\bar{k}$ and $\bar{E}$ are orthogonal.

B.4 Plane Waves in a Good Conductor

In a good conductor, $\sigma/(\omega\varepsilon) \gg 1$ and the propagation constant k_c is approximated as

$$k_c \approx \omega\sqrt{\mu_o\varepsilon}\sqrt{-j\frac{\sigma}{\omega\varepsilon}} = \sqrt{\omega\mu_o\sigma}\left(\frac{1-j}{\sqrt{2}}\right) = \beta - j\alpha \tag{B.19}$$

using the fact that $\sqrt{-j} = (1-j)/\sqrt{2}$. There are some details in this derivation that need attention. The complex number $-j$ in polar form is

$$-j = e^{-j\pi/2} \tag{B.20}$$

and its square root is

$$\sqrt{e^{-j\pi/2}} = \pm e^{-j\pi/4} = \pm\left(1-j\right)/\sqrt{2} \tag{B.21}$$

Which sign should we choose in this case? In this case, we must choose the sign of the imaginary portion to be negative such that the wave attenuates. Therefore, the positive root is chosen.

B.5 Power Flow in Plane Waves

B.5.1 In a Lossless Medium

Assuming propagation in the z-direction, the complex Poynting vector can be written as

$$\begin{aligned}\bar{S}_c = \bar{E}\times\bar{H}^* &= \bar{E}_o e^{-jkz}\times\left(\bar{H}_o e^{-jkz}\right)^* = \bar{E}_o e^{-jkz}\times\left(\bar{H}_o^* e^{+jkz}\right)\\ &= \bar{E}_o\times\bar{H}_o^*\end{aligned} \tag{B.22}$$

In addition, we know that $\bar{H} = (1/\eta)(\hat{k}\times\bar{E})$ where $\hat{k} = \hat{z}$ for propagation in the z-direction. However, we can factor out e^{-jkz} from each side to obtain $\bar{H}_o = (1/\eta)(\hat{z}\times\bar{E}_o)$. Therefore, the complex Poynting vector can be expressed as

$$\bar{S}_c = \bar{E}_o\times\bar{H}_o^* = \left[\bar{E}_o\times\left(\frac{\hat{z}\times\bar{E}_o}{\eta}\right)^*\right] = \frac{1}{\eta^*}[\bar{E}_o\times\left(\hat{z}^*\times\bar{E}_o^*\right)] \tag{B.23}$$

For a lossless material the wave impedance is real and $\eta^* = \eta$. Also, $\hat{z}^* = \hat{z}$ so the complex Poynting vector becomes

$$\bar{S}_c = \frac{1}{\eta}[\bar{E}_o\times\left(\hat{z}\times\bar{E}_o^*\right)] \tag{B.24}$$

Using the triple-product identity $\bar{A}\times(\bar{B}\times\bar{C}) = \bar{B}(\bar{A}\cdot\bar{C}) - \bar{C}(\bar{A}\cdot\bar{B})$, we can write

$$\bar{S}_c = \frac{1}{\eta}[\hat{z}\left(\bar{E}_o\cdot\bar{E}_o^*\right) - \bar{E}_o^*(\bar{E}_o\cdot\hat{z})] \tag{B.25}$$

Since the electric field must be orthogonal to the direction of propagation, $\bar{E}_o\cdot\hat{z} = 0$. This expression then simplifies to

$$\bar{S}_c = \frac{1}{\eta}\left(\bar{E}_o\cdot\bar{E}_o^*\right)\hat{z} = \frac{|\bar{E}_o|^2}{\eta}\hat{z} \tag{B.26}$$

The time-averaged power flow is then

$$\left\langle\bar{S}\right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{S}_c\right\} = \frac{1}{2}\mathcal{R}e\left\{\frac{|\bar{E}_o|^2}{\eta}\hat{z}\right\} = \frac{|\bar{E}_o|^2}{2\eta}\hat{z} \tag{B.27}$$

B.5.2 In a Lossy Medium

In a lossy medium, the time-average is still given by $\langle \mathcal{\bar{S}} \rangle = \langle \bar{\mathcal{E}} \times \bar{\mathcal{H}} \rangle = \frac{1}{2}\mathcal{R}e\{\bar{S}_c\}$. Assuming propagation in the z-direction, the complex Poynting vector is:

$$\bar{S}_c = \bar{E} \times \bar{H}^* = \bar{E}_o e^{-jk_c z} \times \left(\bar{H}_o e^{-jk_c z}\right)^* \tag{B.28}$$

We can write the complex propagation constant as $k_c = \beta - j\alpha$, so therefore we obtain

$$\begin{aligned} \bar{S}_c &= \left(\bar{E}_o e^{-\alpha z} e^{-j\beta z}\right) \times \left(\bar{H}_o e^{-\alpha z} e^{-j\beta z}\right)^* \\ &= \left(\bar{E}_o e^{-\alpha z} e^{-j\beta z}\right) \times \left(\bar{H}_o^* e^{-\alpha z} e^{+j\beta z}\right) \\ &= e^{-2\alpha z}\left(\bar{E}_o \times \bar{H}_o^*\right) \end{aligned} \tag{B.29}$$

In a lossy material, the electric and magnetic fields are related by $\bar{H} = (1/\eta_c)(\hat{z} \times \bar{E})$ for propagation in the z-direction, where η_c is the complex wave impedance. But, we can factor out $e^{-jk_c z}$ from each side to obtain $\bar{H}_o = (1/\eta_c)(\hat{z} \times \bar{E}_o)$. Therefore, the complex Poynting vector can be expressed as

$$\bar{S}_c = e^{-2\alpha z}(\bar{E}_o \times \bar{H}_o^*) = e^{-2\alpha z}[\bar{E}_o \times \left(\frac{\hat{z} \times \bar{E}_o}{\eta_c}\right)^*] = \frac{e^{-2\alpha z}}{\eta_c^*}[\bar{E}_o \times \left(\hat{z} \times \bar{E}_o^*\right)] \tag{B.30}$$

Using the triple-product identity $\bar{A} \times (\bar{B} \times \bar{C}) = \bar{B}(\bar{A} \cdot \bar{C}) - \bar{C}(\bar{A} \cdot \bar{B})$, we can write

$$\bar{S}_c = \frac{e^{-2\alpha z}}{\eta_c^*}[\hat{z}\left(\bar{E}_o \cdot \bar{E}_o^*\right) - \bar{E}_o^*(\bar{E}_o \cdot \hat{z})] \tag{B.31}$$

Since the electric field must be orthogonal to the direction of propagation, $\bar{E}_o \cdot \hat{z} = 0$, and the expression simplifies to

$$\bar{S}_c = \frac{e^{-2\alpha z}}{\eta_c^*}\left(\bar{E}_o \cdot \bar{E}_o^*\right)\hat{z} = \frac{e^{-2\alpha z}|\bar{E}_o|^2}{\eta_c^*}\hat{z} \tag{B.32}$$

The time average power flow is then:

$$\left\langle \bar{\mathcal{S}} \right\rangle = \frac{1}{2}\mathcal{R}e\left\{\bar{S}_c\right\} = \frac{e^{-2\alpha z}|\bar{E}_o|^2}{2}\mathcal{R}e\left\{\frac{1}{\eta_c^*}\right\} \quad \hat{z} \tag{B.33}$$

Letting $\eta_c = |\eta_c|e^{+j\phi}$, this can be rewritten as

$$\begin{aligned} \left\langle \bar{\mathcal{S}} \right\rangle &= \frac{e^{-2\alpha z}|\bar{E}_o|^2}{2}\mathcal{R}e\left\{\frac{1}{|\eta_c|\, e^{-j\phi}}\right\}\hat{z} = \frac{e^{-2\alpha z}|\bar{E}_o|^2}{2\,|\eta_c|}\mathcal{R}e\left\{\cos(\phi) + j\sin(\phi)\right\}\hat{z} \\ &= \frac{e^{-2\alpha z}|\bar{E}_o|^2}{2\,|\eta_c|}\cos(\phi)\hat{z} \end{aligned} \tag{B.34}$$

Appendix C

Derivations for Chapter 3

C.1 Reflection Coefficient from a Dielectric, Normal Incidence

Applying the boundary condition for the tangential electric field yields

$$E_i + E_r = E_t \tag{C.1}$$

while applying the boundary condition for the tangential magnetic fields gives

$$-\frac{E_i}{\eta_o} + \frac{E_r}{\eta_o} = -\frac{E_t}{\eta_2} \tag{C.2}$$

Substituting E_t from Equation C.1 into Equation C.2 yields

$$-\frac{E_i}{\eta_o} + \frac{E_r}{\eta_o} = -\frac{1}{\eta_2}\left(E_i + E_r\right) \tag{C.3}$$

Rearranging, we find

$$\left(\frac{\eta_o}{\eta_2} - 1\right) E_i = \left(-\frac{\eta_o}{\eta_2} - 1\right) E_r \tag{C.4}$$

Solving for the reflection coefficient yields

$$\Gamma = \frac{E_r}{E_i} = \frac{\eta_2 - \eta_o}{\eta_2 + \eta_o} \tag{C.5}$$

To find the transmission coefficient for the electric fields, we use Equation C.1 to find

$$\mathcal{T} = \frac{E_t}{E_i} = \frac{E_r + E_i}{E_i} = \Gamma + 1 \tag{C.6}$$

Then from Equation C.5 we have

$$\Gamma + 1 = \frac{\eta_2 - \eta_o}{\eta_2 + \eta_o} + \frac{\eta_2 + \eta_o}{\eta_2 + \eta_o} = \frac{2\eta_2}{\eta_2 + \eta_o} \tag{C.7}$$

C.2 Surface Current at the Interface of a Perfect Conductor

From the text, the total current per unit width near the surface of a good conductor is

$$\bar{J}_{\text{tot}} = \frac{\sigma_2 \mathcal{T} E_i}{jk_c}\hat{y} = \frac{\sigma_2 E_i}{jk_c}\frac{2\eta_c}{\eta_o + \eta_c}\hat{y} \tag{C.8}$$

But, recall that, for a good conductor,

$$k_c \approx 1/\delta + j/\delta \qquad \text{and} \qquad \eta_c \approx (1+j)/\delta\sigma_2 \tag{C.9}$$

Therefore,

$$\bar{J}_s = \lim_{\sigma_2\to\infty} \bar{J}_{\text{tot}} = \lim_{\sigma_2\to\infty} \frac{\sigma_2 E_i}{\frac{1}{\delta} + j\frac{1}{\delta}} \frac{2\frac{1+j}{\delta\sigma_2}}{\eta_o + \frac{1+j}{\delta\sigma_2}}\hat{y} = \lim_{\sigma_2\to\infty} \frac{E_i}{\frac{1+j}{\delta}} \frac{2\frac{1+j}{\delta}}{\eta_o + \frac{1+j}{\delta\sigma_2}} \quad \hat{y}$$

$$= \lim_{\sigma_2\to\infty} \frac{2E_i}{\eta_o + \frac{1+j}{\delta\sigma_2}}\hat{y} = \lim_{\sigma_2\to\infty} \frac{2E_i}{\eta_o + \sqrt{\frac{\omega\mu_o\sigma_2}{2}}\frac{1+j}{\sigma_2}}\hat{y}$$

$$= \frac{2E_i}{\eta_o}\hat{y} \tag{C.10}$$

C.3 TM Reflection and Transmission Coefficients

C.3.1 Derivation of Coefficients in Terms of Electric Field

For the TM case, the electric field reflection coefficients Γ_{TM} and $\mathcal{T}_{\text{TM}}$ are related to the magnetic field coefficients r_{12} and t_{12}, in the following way. Continuing the notation we have been using in Chapter 3, the incident, reflected, and transmitted magnetic fields are

$$\bar{H}_{1+} = A_1 e^{-j\beta_1 z} e^{-j\kappa x} \hat{y} = H_i e^{-j\beta_1 z} e^{-j\kappa x} \hat{y} \tag{C.11}$$

$$\bar{H}_{1-} = B_1 e^{+j\beta_1 z} e^{-j\kappa x} \hat{y} = H_r e^{+j\beta_1 z} e^{-j\kappa x} \hat{y} \tag{C.12}$$

$$\bar{H}_{2+} = A_2 e^{-j\beta_2 z} e^{-j\kappa x} \hat{y} = H_t e^{-j\beta_2 z} e^{-j\kappa x} \hat{y} \tag{C.13}$$

The incident, reflected, and transmitted electric fields are

$$\begin{aligned} \bar{E}_{1+} &= A_1 \left(\frac{\beta_1}{\omega\varepsilon_1} \hat{x} - \frac{\kappa}{\omega\varepsilon_1} \hat{z} \right) e^{-j\beta_1 z} e^{-j\kappa x} \\ &= A_1 \eta_1 \left(\frac{\beta_1}{k_1} \hat{x} - \frac{\kappa}{k_1} \hat{z} \right) e^{-j\beta_1 z} e^{-j\kappa x} \end{aligned} \tag{C.14}$$

$$\begin{aligned} \bar{E}_{1-} &= B_1 \left(-\frac{\beta_1}{\omega\varepsilon_1} \hat{x} - \frac{\kappa}{\omega\varepsilon_1} \hat{z} \right) e^{+j\beta_1 z} e^{-j\kappa x} \\ &= B_1 \eta_1 \left(-\frac{\beta_1}{k_1} \hat{x} - \frac{\kappa}{k_1} \hat{z} \right) e^{+j\beta_1 z} e^{-j\kappa x} \end{aligned} \tag{C.15}$$

$$\begin{aligned} \bar{E}_{2+} &= A_2 \left(\frac{\beta_2}{\omega\varepsilon_2} \hat{x} - \frac{\kappa}{\omega\varepsilon_2} \hat{z} \right) e^{-j\beta_2 z} e^{-j\kappa x} \\ &= A_2 \eta_2 \left(\frac{\beta_2}{k_2} \hat{x} - \frac{\kappa}{k_2} \hat{z} \right) e^{-j\beta_2 z} e^{-j\kappa x} \end{aligned} \tag{C.16}$$

In deriving the above expressions, we have used the fact that $k/\omega\varepsilon = \omega\sqrt{\mu_o\varepsilon}/\omega\varepsilon = \eta$, and have written the fields such that in the final expressions, the vector in parenthesis is a unit vector pointing in the direction of the field vector (check for yourself that the magnitude is unity). Writing the fields this way reveals that the amplitudes of the incident, reflected, and transmitted electric fields are $E_i = A_1\eta_1$, $E_r = B_1\eta_1$, and $E_t = A_2\eta_2$, respectively. Therefore, the reflection and transmission coefficients in terms of the electric field are:

$$\Gamma_{\text{TM}} = \frac{E_r}{E_i} = \frac{B_1\eta_1}{A_1\eta_1} = r_{12} \tag{C.17}$$

$$\mathcal{T}_{\text{TM}} = \frac{E_t}{E_i} = \frac{A_2\eta_2}{A_1\eta_1} = \frac{\eta_2}{\eta_1} t_{12} = \frac{n_1}{n_2} t_{12} \tag{C.18}$$

C.3.2 Alternate Derivation of Reflection/Transmission Coefficients

It is instructive to derive the expression for Γ_{TM} and $\mathcal{T}_{\mathrm{TM}}$ by starting with the electric fields:

$$\bar{E}_1 = \left[A_1^e e^{-j\beta_1 z}\hat{u}_+ + B_1^e e^{+j\beta_1 z}\hat{u}_-\right] e^{-j\kappa x} \tag{C.19}$$

and

$$\bar{E}_2 = \left[A_2^e e^{-j\beta_2 z}\hat{u}_+ + B_2^e e^{+j\beta_2 z}\hat{u}_-\right] e^{-j\kappa x} \tag{C.20}$$

where $\hat{u}_+$ is a unit vector that is perpendicular the incident wave vector, $\bar{k}_{\text{incident}} = \kappa\hat{x} + \beta_1\hat{z}$, and $\hat{u}_-$ is a unit vector perpendicular to the reflected wave vector, $\bar{k}_{\text{reflected}} = \kappa\hat{x} - \beta_1\hat{z}$. Now matching the tangential electric field and magnetic fields at the interface (at z = 0) it follows that

$$\begin{pmatrix} A_1^e \\ B_1^e \end{pmatrix} = \begin{pmatrix} \beta_1/n_1 & \beta_1/n_1 \\ n_1 & -n_1 \end{pmatrix}^{-1} \begin{pmatrix} \beta_2/n_2 & \beta_2/n_2 \\ n_2 & -n_2 \end{pmatrix} \begin{pmatrix} A_2^e \\ B_2^e \end{pmatrix} \tag{C.21}$$

This relationship can also be written as

$$\begin{pmatrix} A_1^e \\ B_1^e \end{pmatrix} = \frac{1}{\mathcal{T}_{\mathrm{TM}}} \begin{pmatrix} 1 & \Gamma_{\mathrm{TM}} \\ \Gamma_{\mathrm{TM}} & 1 \end{pmatrix} \begin{pmatrix} A_2^e \\ B_2^e \end{pmatrix} \tag{C.22}$$

where

$$\Gamma_{\mathrm{TM}} = \frac{E_r}{E_i} = \frac{\varepsilon_1\beta_2 - \varepsilon_2\beta_1}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2} \tag{C.23}$$

and

$$\mathcal{T}_{\mathrm{TM}\,|} = \frac{E_t}{E_i} = \frac{2\varepsilon_2\beta_1}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2} \tag{C.24}$$

C.3.3 Alternate Derivation of Coefficient Relationship

The relationship between the TM electric field reflection coefficients, Γ_{TM} and $\mathcal{T}_{\mathrm{TM}}$, can be obtained by matching the tangential electric fields at the interface between the two media. Representing the amplitude of the incident field as E_i, the amplitude of the reflected field as E_r, and the amplitude of the transmitted field as E_t, then matching the tangential field components at the interface yields

$$E_i \cos\theta_1 + E_r \cos\theta_1 = E_t \cos\theta_2 \tag{C.25}$$

Substituting $E_r = \Gamma_{\mathrm{TM}} E_i$ and $E_t = \mathcal{T}_{\mathrm{TM}} E_i$ and rearranging, we find

$$1 + \Gamma_{\mathrm{TM}} = \frac{\cos\theta_2}{\cos\theta_1}\mathcal{T}_{\mathrm{TM}} \tag{C.26}$$

Next, using the trigonometric identity $\sin^2\theta + \cos^2\theta = 1$, Equation C.26 can be rewritten as

$$1 + \Gamma_{\mathrm{TM}} = \sqrt{\frac{1 - \sin^2\theta_2}{1 - \sin^2\theta_1}}\,\mathcal{T}_{\mathrm{TM}} \tag{C.27}$$

From Snell's law, $n_1 \sin\theta_1 = n_2 \sin\theta_2$, which is used to derive the final form

$$1 + \Gamma_{\mathrm{TM}} = \sqrt{\frac{1 - \sin^2\theta_1}{1 - \sin^2\theta_1}}\sqrt{\frac{n_1^2}{n_2^2}}\,\mathcal{T}_{\mathrm{TM}} = \frac{n_1^2}{n_2^2}\mathcal{T}_{\mathrm{TM}} \tag{C.28}$$

C.4 Derivation of Brewster's Angle

From the condition for Brewster's angle we have $\varepsilon_2\beta_1 = \varepsilon_1\beta_2$ which can be rewritten in terms of the tangential wavevector component as:

$$\varepsilon_2\sqrt{k_o^2\varepsilon_1 - \kappa^2} = \varepsilon_1\sqrt{k_o^2\varepsilon_2 - \kappa^2} \tag{C.29}$$

Squaring both sides and rearranging we obtain

$$\left(\varepsilon_2^2\varepsilon_1 - \varepsilon_1^2\varepsilon_2\right)k_o^2 = \left(\varepsilon_2^2 - \varepsilon_1^2\right)\kappa^2 \tag{C.30}$$

Solving for κ we obtain

$$\kappa = k_o\sqrt{\frac{\varepsilon_2^2\varepsilon_1 - \varepsilon_1^2\varepsilon_2}{\varepsilon_2^2 - \varepsilon_1^2}} = k_o\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \tag{C.31}$$

We can write the tangential component in terms of the incident angle, θ_B, as

$$\kappa = k_o\sqrt{\varepsilon_1}\sin\theta_B \tag{C.32}$$

Equating the right-hand sides of Equations C.31 and C.32 and solving for $\sin\theta_B$ we find

$$\sin\theta_B = \sqrt{\frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \tag{C.33}$$

Since $\sin^2\theta + \cos^2\theta = 1$, we can also obtain the following:

$$\cos\theta_B = \sqrt{1 - \sin^2\theta_B} = \sqrt{1 - \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2}} = \sqrt{\frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2}} \tag{C.34}$$

Using Equations C.33 and C.34, we find

$$\tan\theta_B = \frac{\sin\theta_B}{\cos\theta_B} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} \tag{C.35}$$

C.5 Power Conservation at an Interface (TE Case)

Equating the time-average, z-directed power at the interface yields

$$\beta_1 - \beta_1|\Gamma_{\mathrm{TE}}|^2 = \mathcal{R}e\left\{\beta_2^*\right\}|\mathcal{T}_{\mathrm{TE}}|^2 \tag{C.36}$$

Substituting the expressions for Γ_{TE} and $\mathcal{T}_{\mathrm{TE}}$ yields

$$\beta_1 - \beta_1\left|\frac{\beta_1 - \beta_2}{\beta_1 + \beta_2}\right|^2 = \mathcal{R}e\left\{\beta_2^*\right\}\left|\frac{2\beta_1}{\beta_1 + \beta_2}\right|^2 \tag{C.37}$$

Since $|x|^2 = xx^*$ and $\beta_1^* = \beta_1$, Equation C.37 can be rewritten as

$$1 - \left(\frac{\beta_1 - \beta_2}{\beta_1 + \beta_2}\right)\left(\frac{\beta_1 - \beta_2^*}{\beta_1 + \beta_2^*}\right) = \frac{\mathcal{R}e\left\{\beta_2^*\right\}}{\beta_1}\left(\frac{2\beta_1}{\beta_1 + \beta_2}\right)\left(\frac{2\beta_1}{\beta_1 + \beta_2^*}\right) \tag{C.38}$$

Carrying out the multiplication and simplifying yields

$$\beta_1^2 + \beta_1\beta_2 + \beta_1\beta_2^* + \beta_2\beta_2^* - \left(\beta_1^2 - \beta_1\beta_2 - \beta_1\beta_2^* + \beta_2\beta_2^*\right) = 4\beta_1\mathcal{R}e\left\{\beta_2^*\right\} \tag{C.39}$$

which simplifies further to

$$2\beta_1(\beta_2 + \beta_2^*) = 4\beta_1\mathcal{R}e\left\{\beta_2^*\right\} \tag{C.40}$$

Note that $\beta_2 + \beta_2^* = 2\mathcal{R}e\{\beta_2^*\}$, so that power is conserved across the interface.

C.6 Power Conservation at an Interface (TM Case)

Equating the time-average, z-directed power at the interface yields

$$\beta_1 - \beta_1 |\Gamma_{\mathrm{TM}}|^2 = \mathcal{R}e\left\{\beta_2^*\right\} |\mathcal{T}_{\mathrm{TM}}|^2 \tag{C.41}$$

Substituting the expressions for Γ_{TM} and $\mathcal{T}_{\mathrm{TM}}$ yields

$$\beta_1 - \beta_1 \left|\frac{\varepsilon_2\beta_1 - \varepsilon_1\beta_2}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2}\right|^2 = \mathcal{R}e\left\{\beta_2\right\} \left|\sqrt{\frac{\varepsilon_1}{\varepsilon_2}} \frac{2\varepsilon_2\beta_1}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2}\right|^2 \tag{C.42}$$

Since $|x|^2 = xx^*$ and $\beta_1^* = \beta_1$, Equation C.42 can be rewritten as

$$1 - \left(\frac{\varepsilon_2\beta_1 - \varepsilon_1\beta_2}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2}\right)\left(\frac{\varepsilon_2\beta_1 - \varepsilon_1\beta_2^*}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2^*}\right)$$

$$= \frac{\mathcal{R}e\left\{\beta_2\right\}}{\beta_1}\left(\frac{\varepsilon_1}{\varepsilon_2}\right)\left(\frac{2\varepsilon_2\beta_1}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2}\right)\left(\frac{2\varepsilon_2\beta_1}{\varepsilon_2\beta_1 + \varepsilon_1\beta_2^*}\right) \tag{C.43}$$

Carrying out the multiplication and simplifying yields

$$\varepsilon_2^2\beta_1^2 + \varepsilon_1\varepsilon_2(\beta_1\beta_2 + \beta_1\beta_2^*) + \varepsilon_1^2\beta_2\beta_2^* - \left[\varepsilon_2^2\beta_1^2 - \varepsilon_1\varepsilon_2(\beta_1\beta_2 + \beta_1\beta_2^*) + \varepsilon_1^2\beta_2\beta_2^*\right]$$

$$= 4\varepsilon_1\varepsilon_2\beta_1\mathcal{R}e\left\{\beta_2\right\} \tag{C.44}$$

which simplifies further to

$$2\varepsilon_1\varepsilon_2\beta_1(\beta_2 + \beta_2^*) = 4\varepsilon_1\varepsilon_2\beta_1\mathcal{R}e\left\{\beta_2\right\} \tag{C.45}$$

Note that $\beta_2 + \beta_2^* = 2\mathcal{R}e\{\beta_2\}$, so that power is conserved across the interface.

Appendix D

Derivations for Chapter 4

D.1 Propagation Constant of a Parallel Plate Waveguide

The propagation constant of the parallel plate waveguide can be written in terms of the cutoff frequency as follows. The propagation constant for the m^{th} mode of the parallel plate waveguide is

$$\beta_m = \sqrt{k_1^2 - \kappa^2} = \sqrt{k_1^2 - (m\pi/a)^2} = k_1\sqrt{1 - (1/k_1)^2\,(m\pi/a)^2}$$

$$= k_1\sqrt{1 - [m\pi c/\,(a2\pi f\sqrt{\varepsilon_r})]^2} = k_1\sqrt{1 - [mc/\,(2a\sqrt{\varepsilon_r}f)]^2}$$

$$= k_1\sqrt{1 - f_{c,m}^2/f^2} \tag{D.1}$$

And since $f = c/\lambda$, it follows that

$$f_{c,m}^2/f^2 = \left(c/\lambda_{c,m}^2\right)^2/\,(c/\lambda)^2 = \lambda^2/\lambda_{c,m}^2 \tag{D.2}$$

therefore the propagation constant can be rewritten as

$$\beta_m = k_1\sqrt{1 - f_{c,m}^2/f^2} = k_1\sqrt{1 - \lambda^2/\lambda_{c,m}^2} \tag{D.3}$$

D.2 Guide Wavelength in a Parallel Plate Waveguide

Using the results from the previous section, the wavelength in the z-direction is

$$\lambda_z = \frac{2\pi}{\beta_m} = \frac{2\pi}{k_1\sqrt{1 - f_{c,m}^2/f^2}}$$

$$= \frac{2\pi\lambda}{2\pi\sqrt{\varepsilon_r}\sqrt{1 - f_{c,m}^2/f^2}} = \frac{\lambda/\sqrt{\varepsilon_r}}{\sqrt{1 - f_{c,m}^2/f^2}} \tag{D.4}$$

D.3 Phase Velocity in a Parallel Plate Waveguide

The phase velocity can also be rewritten in terms of the cutoff frequency/wavelength as follows:

$$v_p = \frac{\omega}{\beta_m} = \frac{\omega}{k_1\sqrt{1 - f_{c,m}^2/f^2}} = \frac{\omega c}{\omega\sqrt{\varepsilon_r}\sqrt{1 - f_{c,m}^2/f^2}}$$

$$= \frac{c/\sqrt{\varepsilon_r}}{\sqrt{1 - f_{c,m}^2/f^2}} \tag{D.5}$$

Using the result from Equation D.2, the phase velocity can be written as

$$v_p = \frac{\omega}{\beta_m} = \frac{c/\sqrt{\varepsilon_r}}{\sqrt{1 - f_{c,m}^2/f^2}} = \frac{c/\sqrt{\varepsilon_r}}{\sqrt{1 - \lambda^2/\lambda_{c,m}^2}} \tag{D.6}$$

D.4 TE Helmholtz Equation in a Waveguide

The Helmholtz equation for the TE case in a dielectric waveguide is derived as follows. First, substitute the field solution $\bar{E} = E_y(x)e^{-j\beta z}\hat{y}$ into the wave equation $\nabla^2\bar{E} + k^2\bar{E} = 0$. This yields

$$\nabla^2\left[E_y(x)e^{-j\beta z}\right] + k^2[E_y(x)e^{-j\beta z}] = 0 \tag{D.7}$$

Writing out the Laplacian operator we find

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\left(E_y e^{-j\beta z}\right) + k^2\left(E_y e^{-j\beta z}\right)$$

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\left(E_y e^{-j\beta z}\right) + (-j\beta)^2\left(E_y e^{-j\beta z}\right) + k^2\left(E_y e^{-j\beta z}\right) =$$

$$= \nabla_t^2\left(E_y e^{-j\beta z}\right) + \left(k^2 - \beta^2\right)\left(E_y e^{-j\beta z}\right) \tag{D.8}$$

This is written compactly as

$$\nabla_t^2\bar{E} + \left(k^2 - \beta^2\right)\bar{E} = 0 \tag{D.9}$$

D.5 Alternative TE Field Derivation in a Planar Waveguide

The field solutions for the TE modes of a dielectric planar waveguide may be obtained in an alternative manner by returning to the bouncing wave picture shown in Figure D.1. In each region, the bouncing upward and downward waves are described by the equations

$$\begin{aligned} \bar{E}_u &= Ae^{-j\kappa x}e^{-j\beta z} \qquad \text{(upward wave)} \\ \bar{E}_d &= Ae^{+j\kappa x}e^{-j\beta z} \qquad \text{(downward wave)} \end{aligned} \tag{D.10}$$

where A is an arbitrary constant. In the upper cladding region, there is only an upward travelling wave

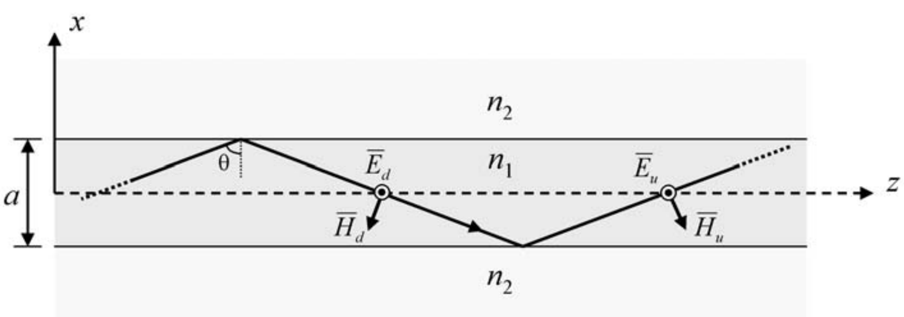


Figure D.1: "Bouncing wave" picture of waveguiding in a planar dielectric waveguide. The upwards and downwards fields are indicated by the subscripts "u" and "d", respectively.

with $\kappa_2 = -j\gamma_2$ which leads to a decaying exponential starting at $x = a/2$.

$$\bar{E} = E_o e^{-\gamma_2(x-a/2)}e^{-j\beta z}\ \hat{y} \qquad (x \geq a/2) \tag{D.11}$$

Similarly, in the lower cladding there is only a downward wave with $\kappa_2 = -j\gamma_2$, which leads to

$$\bar{E} = E_o e^{+\gamma_2(x+a/2)}e^{-j\beta z}\hat{y} \qquad (x \leq -a/2) \tag{D.12}$$

In the core, both an upward and downward traveling wave exist. Taking $x = 0$ as a reference point, the difference in phase, $\Delta\varphi$, between the upward and downward travelling waves at $x = 0$ should be one-half of the total round-trip accumulated phase, which is $2\pi\nu$. The phase difference at $x = 0$ can then be written as $\Delta\varphi = 2\pi m$ for even values of ν, or $\Delta\varphi = (2m + 1)\pi$ for odd values of ν, where $m = 0, 1, 2, 3....$ In order to retain the symmetry about $x = 0$, we apply half of the phase difference to the upward wave and half of the phase difference to the downward wave. Therefore, for *even values* of ν the total field in the core is

$$\bar{E}_{\text{even}} = \bar{E}_u e^{-j\Delta\varphi/2} + \bar{E}_d e^{+j\Delta\varphi/2}$$

$$= Ae^{-j\kappa x} e^{-j\beta z} e^{-j\pi m} + Ae^{+j\kappa x} e^{-j\beta z} e^{+j\pi m}$$

$$= \pm \left(Ae^{-j\kappa x} + Ae^{+j\kappa x}\right) e^{-j\beta z} = \pm 2A \cos(\kappa x) e^{-j\beta z} \tag{D.13}$$

For *odd values* of ν the total field in the core is

$$\bar{E}_{\text{odd}} = \bar{E}_u e^{-j\Delta\varphi/2} + \bar{E}_d e^{+j\Delta\varphi/2}$$

$$= Ae^{-j\kappa x} e^{-j\beta z} e^{-j\pi m} e^{-j\pi/2} + Ae^{+j\kappa x} e^{-j\beta z} e^{+j\pi m} e^{+j\pi/2}$$

$$= \pm \left(jAe^{-j\kappa x} - jAe^{+j\kappa x}\right) e^{-j\beta z} = \pm 2A \sin(\kappa x) e^{-j\beta z} \tag{D.14}$$

These are exactly the field solutions for the core region given in the text, taking into consideration the arbitrary normalization constant $\pm 2A$.

Appendix E

Derivations for Chapter 5

E.1 Alternate Derivation of Group Velocity in a Plasma

An alternate method for deriving the group velocity is as follows. Begin with the propagation constant of a wave in a plasma

$$k_p = \frac{1}{c}\sqrt{\omega^2 - \omega_p^2} \tag{E.1}$$

Solve the expression for ω

$$\omega = \sqrt{c^2 k_p^2 + \omega_p^2} \tag{E.2}$$

Differentiate Equation E.2 with respect to k_p:

$$v_g = \frac{d\omega}{dk_p} = \frac{1}{2}\left(c^2 k_p^2 + \omega_p^2\right)^{-1/2}\left(2c^2 k_p\right) = c\sqrt{\frac{c^2 k_p^2}{c^2 k_p^2 + \omega_p^2}} \tag{E.3}$$

Then simplify this result by substituting k_p from Equation E.1 to obtain

$$v_g = c\sqrt{\frac{\omega^2 - \omega_p^2}{\left(\omega^2 - \omega_p^2\right) + \omega_p^2}} = c\sqrt{1 - \frac{\omega_p^2}{\omega^2}} \tag{E.4}$$

E.2 Group Velocity in a Parallel Plate Waveguide

We begin with the expression for the propagation constant derived in Chapter 4, and solve for ω:

$$\omega = \frac{c}{\sqrt{\varepsilon_r}}\sqrt{\beta_m^2 + (m\pi/a)^2} \tag{E.5}$$

Assuming that ε_r is not a function of frequency, the group velocity of the m^{th} mode is then

$$v_{g,m} = \frac{d\omega}{d\beta_m} = \frac{c}{\sqrt{\varepsilon_r}}\frac{1}{2}\left\{\beta_m^2 + (m\pi/a)^2\right\}^{-1/2}(2\beta_m) = \frac{c}{\sqrt{\varepsilon_r}}\frac{\beta_m}{\sqrt{\beta_m^2 + (m\pi/a)^2}} \tag{E.6}$$

Now we use the fact that $\beta_m^2 + (m\pi/a)^2 = \beta_m^2 + \kappa_1^2 = k_1^2$ to write

$$v_{g,m} = \frac{c}{\sqrt{\varepsilon_r}}\frac{\beta_m}{\sqrt{k_1^2}} = \frac{c\beta_m}{\sqrt{\varepsilon_r}k_1} \tag{E.7}$$

where $k_1 = \omega\sqrt{\mu_o\varepsilon}$. Also in Chapter 4 we found that $\beta_m = k_1\sqrt{1 - f_{c,m}^2/f^2}$ (see Appendix D), so this leads to

$$v_{g,m} = \frac{ck_1\sqrt{1 - f_{c,m}^2/f^2}}{\sqrt{\varepsilon_r}k_1} = \frac{c}{\sqrt{\varepsilon_r}}\sqrt{1 - f_{c,m}^2/f^2} \tag{E.8}$$

Appendix F

Derivations for Chapter 6

F.1 Approximation of $1/R_+$

In order to determine an expression for the dipole moment of two charges, we must develop an approximate expression for the distance to the charges. As an example, we consider the approximation of R_+, the distance from the observation point to the positive charge.

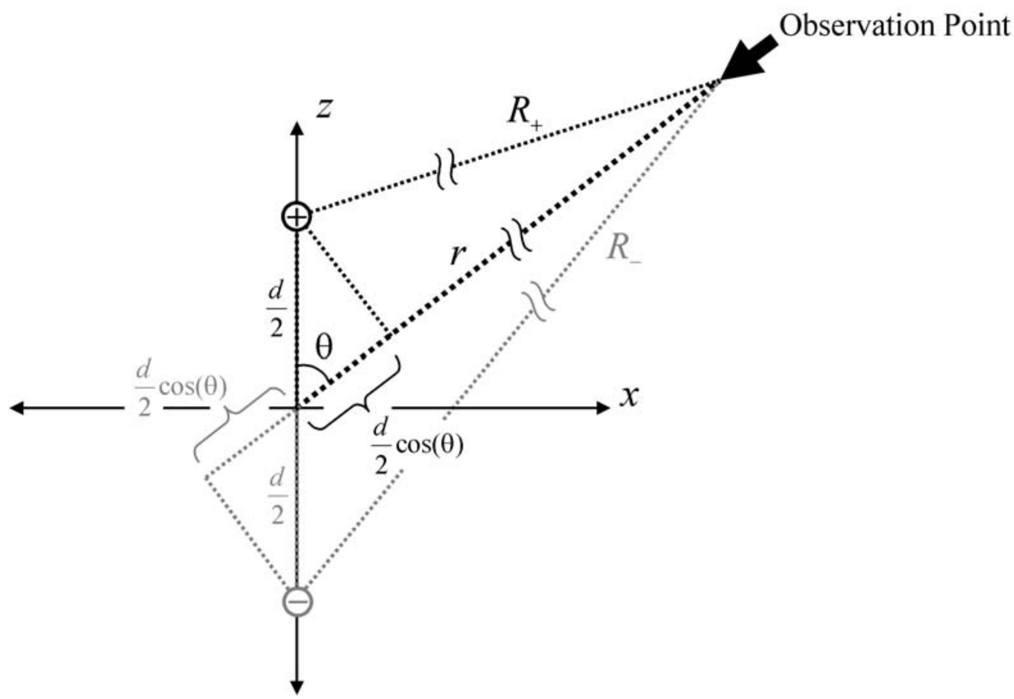


Figure F.1: Two static charges separated by a distance d. The separation between the charges is assumed to be much larger than the distance to the observation point.

From Figure F.1, we see that

$$R_+ \approx \left(r - \frac{d}{2}\cos\theta\right) \tag{F.1}$$

In our expressions for the scalar potential, we require the inverse of R_+, i.e.

$$\frac{1}{R_+} \approx \left(r - \frac{d}{2}\cos\theta\right)^{-1} = \frac{1}{r}\left(1 - \frac{d}{2r}\cos\theta\right)^{-1} \tag{F.2}$$

Assuming that $d \ll r$, then $\frac{d}{2r}\cos\theta \ll 1$. Defining $\delta = \frac{d}{2r}\cos\theta$, we can use the power series expansion:

$$(1-\delta)^{-1} = 1 + \delta + \delta^2 + \delta^3 + ... \approx 1 + \delta \tag{F.3}$$

which yields the approximation

$$\frac{1}{R_+} \approx \frac{1}{r}(1+\delta) = \frac{1}{r}\left(1 + \frac{d}{2r}\cos\theta\right) \tag{F.4}$$

A similar approach is used to determine an approximation for $1/R\_$.

F.2 Inhomogeneous Vector Wave Equation

Substitution of the vector potential $\nabla \times \bar{\mathcal{A}} = \mu_o \bar{\mathcal{H}}(\bar{r}, t)$ into Ampere's law yields

$$\nabla \times \left(\nabla \times \bar{\mathcal{A}}\right) = \mu_o \varepsilon_o \frac{\partial \bar{\mathcal{E}}}{\partial t} + \mu_o \bar{\mathcal{J}} \tag{F.5}$$

We have previously defined

$$-\nabla \Phi(\bar{r}, t) = \bar{\mathcal{E}}(\bar{r}, t) + \frac{\partial \bar{\mathcal{A}}(\bar{r}, t)}{\partial t} \tag{F.6}$$

We rearrange Equation F.6 to solve for $\bar{\mathcal{E}}$ and substitute into Equation F.5 to yield

$$\nabla \times \left(\nabla \times \bar{\mathcal{A}}\right) = \mu_o \varepsilon_o \frac{\partial}{\partial t}\left(-\nabla \Phi - \frac{\partial \bar{\mathcal{A}}}{\partial t}\right) + \mu_o \bar{\mathcal{J}}$$

$$= \mu_o \bar{\mathcal{J}} - \mu_o \varepsilon_o \nabla \frac{\partial \Phi}{\partial t} - \mu_o \varepsilon_o \frac{\partial^2 \bar{\mathcal{A}}}{\partial t^2} \tag{F.7}$$

Using the vector identity $\nabla \times (\nabla \times \bar{\mathcal{A}}) = \nabla(\nabla \cdot \bar{\mathcal{A}}) - \nabla^2 \bar{\mathcal{A}}$, we find

$$\nabla\left(\nabla \cdot \bar{\mathcal{A}}\right) - \nabla^2 \bar{\mathcal{A}} = \mu_o \bar{\mathcal{J}}(\bar{r}, t) - \mu_o \varepsilon_o \nabla \frac{\partial \Phi}{\partial t} - \mu_o \varepsilon_o \frac{\partial^2 \bar{\mathcal{A}}}{\partial t^2} \tag{F.8}$$

This is then rearranged to yield

$$\nabla^2 \bar{\mathcal{A}} - \mu_o \varepsilon_o \frac{\partial^2 \bar{\mathcal{A}}}{\partial t^2} = -\mu_o \bar{\mathcal{J}} + \nabla\left(\nabla \cdot \bar{\mathcal{A}} + \mu_o \varepsilon_o \frac{\partial \Phi}{\partial t}\right) \tag{F.9}$$

We are free to choose the divergence of $\bar{\mathcal{A}}$, so for convenience we choose

$$\nabla \cdot \bar{\mathcal{A}} = -\mu_o \varepsilon_o \frac{\partial \Phi}{\partial t} \tag{F.10}$$

which is known as the Lorentz gauge. Applying the Lorentz gauge, Equation F.9 reduces to

$$\nabla^2 \bar{\mathcal{A}} - \mu_o \varepsilon_o \frac{\partial^2 \bar{\mathcal{A}}}{\partial t^2} = -\mu_o \bar{\mathcal{J}} \tag{F.11}$$

which is the *inhomogeneous vector wave equation*.

F.3 Inhomogeneous Scalar Wave Equation

To derive the inhomogeneous scalar wave equation, we rearrange Equation F.6 to solve for $\bar{\mathcal{E}}$ and substitute into Gauss' law to yield

$$\nabla \cdot \left(-\nabla \Phi - \frac{\partial \bar{\mathcal{A}}}{\partial t}\right) = \frac{\rho_v}{\varepsilon_o} \tag{F.12}$$

We can then rewrite the left-hand side as

$$\nabla \cdot \left(-\nabla \Phi - \frac{\partial \bar{\mathcal{A}}}{\partial t}\right) = -\nabla^2 \Phi - \frac{\partial}{\partial t}\left(\nabla \cdot \bar{\mathcal{A}}\right). \tag{F.13}$$

Using the Lorentz gauge ($\nabla \cdot \bar{\mathcal{A}} = -\mu_o \varepsilon_o \partial\Phi/\partial t$), we find

$$\frac{\partial}{\partial t}\left(\nabla \cdot \bar{\mathcal{A}}\right) = -\mu_o \varepsilon_o \frac{\partial^2 \Phi}{\partial t^2} \tag{F.14}$$

Then Equation F.12 reduces to

$$\nabla^2 \Phi - \mu_o \varepsilon_o \frac{\partial^2 \Phi}{\partial t^2} = -\frac{\rho_v}{\varepsilon_o} \tag{F.15}$$

which is the *inhomogeneous scalar wave equation*.

F.4 Scalar Potential of an Isolated Oscillating Charge

Assume an oscillating charge $q(t) = q_o \cos(\omega t)$ located at the origin (in free space). We wish to demonstrate that the following scalar potential

$$\Phi = \frac{q_o}{4\pi\varepsilon_o r} e^{-jk_o r} \tag{F.16}$$

is the solution to the inhomogeneous scalar wave equation (Equation F.15) in phasor form:

$$\nabla^2 \Phi + \omega^2 \mu_o \varepsilon_o \Phi = -\frac{\rho_v}{\varepsilon_o} \tag{F.17}$$

Demonstration of this will take two steps. First, we will show that Equation F.17 is satisfied everywhere *except* the origin. The charge only exists at the origin, so everywhere except the origin Equation F.17 reduces to

$$\nabla^2 \Phi + \omega^2 \mu_o \varepsilon_o \Phi = 0 \tag{F.18}$$

In spherical coordinates, the Laplacian can be written as

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2}{\partial \phi^2} + \frac{\cos\theta}{r^2 \sin\theta}\frac{\partial}{\partial \theta} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \tag{F.19}$$

Evaulating the first term of Equation F.18 yields

$$\nabla^2 \Phi = \left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\left(\frac{q_o}{4\pi\varepsilon_o r} e^{-jk_o r}\right)$$

$$= \frac{q_o}{4\pi\varepsilon_o}\left\{-\frac{k_o^2}{r} + \frac{j2k_o}{r^2} + \frac{2}{r^3} + \frac{2}{r}\left(\frac{-jk_o}{r} - \frac{1}{r^2}\right)\right\} e^{-jk_o r}$$

$$= -\frac{q_o}{4\pi\varepsilon_o}\frac{k_o^2}{r} e^{-jk_o r} \tag{F.20}$$

Equation F.18 then becomes

$$\frac{q_o}{4\pi\varepsilon_o r}\left(\omega^2 \mu_o \varepsilon_o - k_o^2\right) e^{-jk_o r} = 0 \tag{F.21}$$

Which is true since $k_o^2 = \omega^2 \mu_o \varepsilon_o$ in free space.

The next step is to verify that Equation F.17 is satisfied at the origin. We integrate Equation F.17 in the neighborhood of $r = 0$ over a small spherical volume of radius r centered on the origin:

$$\int_V \left(\nabla^2 \Phi + \omega^2 \mu_o \varepsilon_o \Phi\right) dV = -\int_V \left(\frac{\rho_v}{\varepsilon_o}\right) dV \tag{F.22}$$

In the limit where r is small, the integration of the second term on the left-hand side gives no contribution, since

$$\lim_{r\to 0}\int_V \left(\omega^2\mu_o\varepsilon_o\Phi\right)dV = \omega^2\mu_o\varepsilon_o \lim_{r\to 0}\int_V \Phi dV$$

$$= \omega^2\mu_o\varepsilon_o \lim_{r\to 0}\int_V \frac{q_o e^{-jk_o r}}{4\pi\varepsilon_o r}dV = \omega^2\mu_o\varepsilon_o \lim_{r\to 0}\int_0^r \frac{q_o e^{-jk_o r}}{4\pi\varepsilon_o r}\left(4\pi r^2\right)dr$$

$$= \omega^2\mu_o q_o \lim_{r\to 0}\int_0^r r e^{-jk_o r}dr = 0 \tag{F.23}$$

Integration of the first term on the left-hand side is done using Gauss' theorem (refer to Chapter 1) to convert the volume integral to one enclosing a surface:

$$\int_V \nabla^2\Phi dV = \int_V \nabla\cdot\nabla\Phi dV = \oint_S \nabla\Phi\cdot d\bar{S} \tag{F.24}$$

Since the scalar potential varies only in the $\hat{r}$-direction, $\nabla\Phi = \partial\Phi/\partial r$ and the surface integration yields simply the surface area of a sphere of radius r, and therefore

$$\oint_S \nabla\Phi\cdot d\bar{S} = 4\pi r^2\frac{\partial\Phi}{\partial r} = 4\pi r^2\frac{\partial}{\partial r}\left(\frac{q_o}{4\pi\varepsilon_o r}e^{-jk_o r}\right)$$

$$= \frac{q_o}{\varepsilon_o}r^2\frac{\partial}{\partial r}\left(\frac{e^{-jk_o r}}{r}\right) = \frac{q_o}{\varepsilon_o}\left(-jk_o r - 1\right)e^{-jk_o r} \tag{F.25}$$

Since the charge is located at the origin, the volume integral on the right-hand side of Equation F.22 yields

$$-\int_V \left(\frac{\rho_v}{\varepsilon_o}\right)dV = -\frac{q_o}{\varepsilon_o} \tag{F.26}$$

Therefore, as $r \to 0$ Equation F.22 reduces to

$$\frac{q_o}{\varepsilon_o}\left(-jk_o r - 1\right)e^{-jk_o r} = -\frac{q_o}{\varepsilon_o} \tag{F.27}$$

which is indeed satisfied for $r = 0$. We then conclude that the scalar potential given in Equation F.16 is a solution to the inhomogeneous scalar wave equation for an oscillating point charge at the origin.

Appendix G

Glossary of Commonly Used Symbols

Following is a summary of important symbols used throughout the text. Also presented in parentheses is the chapter in which the symbol was first introduced.

a: The width of the guiding region in a parallel plate or dielectric planar waveguide, or the width in the x-direction of a rectangular waveguide (Chapter 4).

$\bar{A}(\bar{r})$: A vector field, which is a field that has a magnitude and direction at every point in space (Chapter 1).

α: The attenuation constant of a wave, which is the imaginary portion of the propagation constant. The attenuation constant has units of Np/m (Chapter 2).

b: The normalized propagation constant in a dielectric planar waveguide, or the width in the y-direction of a rectangular waveguide (Chapter 4).

$\bar{\mathcal{B}}(\bar{r},t), \bar{B}(\bar{r})$: The magnetic flux density in time-domain and phasor form, respectively. The magnetic flux has units of T (Chapter 2).

β: The z-component of the wave vector, with units of rad/m (Chapter 3).

c: The speed of light in a vacuum. $c \approx 2.998 \times 10^8$ m/s (Chapter 2).

$\bar{\mathcal{D}}(\bar{r},t), \bar{D}(\bar{r})$: The electric flux density in time-domain and phasor form, respectively. The electric flux has units of C/m^2 (Chapter 2).

∇: The gradient operator, defined as $\nabla = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}$ (Chapter 1).

δ: The skin depth, which is the depth at which the electric field decreases to e^{-1} of its initial value in a lossy medium (Chapter 2).

$\bar{\mathcal{E}}(\bar{r},t), \bar{E}(\bar{r})$: The electric field intensity in time-domain and phasor form, respectively. The electric field has units of V/m (Chapter 2).

ε: The electric permittivity, with units of F/m. In free space, $\varepsilon_o = 8.8542 \times 10^{-12}$ F/m (Chapter 2).

ε_r: The relative permittivity with respect to free space, defined as $\varepsilon_r = \varepsilon/\varepsilon_o$ (Chapter 2).

f_c: The cutoff frequency of a waveguide mode, which is the lowest frequency at which a particular mode can propagate without loss. The cutoff frequency has units of Hz (Chapter 4).

η: The wave impedance of a medium, with units of Ω. The wave impedance of free space is $\eta_o \approx 120\pi \approx 377$ Ω (Chapter 2).

η_g: The waveguide impedance of a mode, defined with respect to power flow in the z-direction. The waveguide impedance has units of Ω (Chapter 4).

Γ: The electric field reflection coefficient, defined as the ratio of the reflected field amplitude to the incident field amplitude (Chapter 3).

$\Gamma_{a,b}$: The electric field reflection coefficient between two dissimilar waveguide sections (Chapter 4).

$\bar{\mathcal{H}}(\bar{r},t), \bar{H}(\bar{r})$: The magnetic field intensity in time-domain and phasor form, respectively. The magnetic field has units of A/m (Chapter 2).

$\bar{\mathcal{J}}(\bar{r},t), \bar{J}(\bar{r})$: The electric current density in time-domain and phasor form, respectively. The electric current density has units of A/m^2 (Chapter 2).

$\bar{\mathcal{J}}_s(\bar{r},t), \bar{J}_s(\bar{r})$: The surface current at an interface between two media in time-domain and phasor form, respectively. The surface current has units of A/m (Chapter 3).

$\bar{k}$: The wave vector, which points in the direction of propagation and is defined as $\bar{k} = k\hat{k}$ (Chapter 2).

k: The magnitude of the wave vector, with units of rad/m. In free space, $k_o = \omega\sqrt{\mu_o\varepsilon_o} = \omega/c = 2\pi/\lambda_o$ (Chapter 2).

$\hat{k}$: A unit vector in the direction of wave propagation (Chapter 2).

κ: The x-component of the wave vector, with units of rad/m (Chapter 3).

λ: The wavelength of a wave in space, with units of m. The wavelength in free space is signified by the symbol λ_o (Chapter 2).

λ_c: The cutoff wavelength of a waveguide mode, which is the largest wavelength at which a particular mode can propagate without loss. The cutoff wavelength has units of m (Chapter 4).

m: The mode number in a parallel plate waveguide, or the mode number in the x-direction in a rectangular metallic waveguide (Chapter 4).

μ: The magnetic permeability, with units of H/m. In free space, $\mu_o = 4\pi \times 10^{-7}$ H/m (Chapter 2).

n: The index of refraction of a medium. Defined as $n = \sqrt{\varepsilon_r}$ (Chapter 2). Also represents the mode number in the y-direction in a rectangular metallic waveguide (Chapter 4).

$\hat{n}$: A unit vector that is normal to a surface of integration (used in Gauss' Theorem) (Chapter 1).

$\hat{n}_{21}$: A unit vector pointing from region 2 to region 1 at an interface between two media (Chapter 3).

n_{eff}: The effective index of a waveguide mode in a dielectric waveguide. Defined as $n_{\text{eff}} = \beta/k_o$ (Chapter 4).

ν: The mode number in a dielectric planar waveguide (Chapter 4).

$p(\theta,\phi)$: The radiation pattern of an antenna, defined as the gain normalized to have a maximum value of unity (Chapter 6).

$\Phi(\bar{r})$: A scalar field, which is a field that has only a magnitude at every point in space (Chapter 1).

$\bar{r}$: The position vector. In Cartesian coordinates, $\bar{r} = x\hat{x} + y\hat{y} + z\hat{z}$ (Chapter 1).

R: The power reflectance, which is the ratio of reflected to incident power at an interface (Chapter 3).

r_{12}: The reflection coefficient in the matrix formulation. It is with respect to the electric fields for the TE case and with respect to the magnetic fields for the TM case (Chapter 3).

$\rho_s(\bar{r},t), \rho_s(\bar{r})$: The surface charge at the interface between two media in time-domain and phasor form, respectively. The surface charge has units of C/m^2 (Chapter 3).

$\rho_v(\bar{r},t), \rho_v(\bar{r})$: The electric charge density in time-domain and phasor form, respectively. The electric charge density has units of C/m^3 (Chapter 2).

$\langle\bar{\mathcal{S}}\rangle$: The time-average power flow per unit area, with units of W/m^2. Defined as $\langle\bar{\mathcal{S}}\rangle = \langle\bar{\mathcal{E}} \times \bar{\mathcal{H}}\rangle = \frac{1}{2}\mathcal{R}e\{\bar{S}_c\}$ (Chapter 2).

$\bar{S}_c$: The complex Poynting vector. Defined as $\bar{S}_c = \bar{E} \times \bar{H}^*$ (Chapter 2).

σ: The conductivity of a medium, with units of S/m (Chapter 2).

T: The period of a wave in time (Chapter 2). Also represents the power transmittance, which is the ratio of transmitted to incident power at an interface (Chapter 3).

$\mathcal{T}$: The electric field transmission coefficient, defined as the ratio of the transmitted field amplitude to the incident field amplitude (Chapter 3).

t_{12}: The transmission coefficient in the matrix formulation. It is with respect to the electric fields for the TE case and with respect to the magnetic fields for the TM case (Chapter 3).

TE: Transverse electric polarization, which means that the electric field is perpendicular to the plane of incidence. Also called "perpendicular polarization" (Chapter 3).

TM: Transverse magnetic polarization, which means that the magnetic field is perpendicular to the plane of incidence. Also called "parallel polarization" (Chapter 3).

θ_B: Brewster's angle. At θ_B, the TM reflection coefficient is zero (Chapter 3).

θ_c: The critical angle. If $n_1 > n_2$, then beyond θ_c total internal reflection (TIR) occurs (Chapter 3).

V: The normalized frequency (or "V-number") in a dielectric planar waveguide (Chapter 4).

v_g: The group velocity, which is the velocity at which the envelope of an electromagnetic pulse travels. The group velocity has units of m/s (Chapter 5).

v_p: The phase velocity, which is the velocity at which a particular point on a wave moves. The phase velocity has units of m/s (Chapter 5).

Bibliography

[1] H. H. Skilling. *Fundamentals of Electrical Waves.* Wiley and Sons, Second Edition, 1948

[2] G. Schmid, G. Neubauer, P. R. Mazal, F. Alesch, and U. M. Illievich, "Dielectric properties of brain tissue: measurements on humans," presented at the 5th COST 281 MCM and Workshop, Budapest, Nov. 15th and 16th, 2003.

[3] S. Ramo, J. R. Whinnery, and T. Van Duzer. *Fields and Waves in Communication Electronics.* Wiley and Sons, Third Edition, 1994

[4] http://www.teraspan.com

[5] http://www.accesscomms.com.au

[6] http://space-machine.com

[7] http://www.corningcablesystems.com

[8] http://www.ieee.ca/millennium/radio/radio_differences.html

[9] http://www.ieee.org/organizations/history_center/legacies/tesla.html