Electromagnetic Theory

Chapter · October 2015

DOI: 10.1093/acprof:oso/9780198738770.003.0002

Bob D Guenther
Duke University

Electromagnetic Theory

2

2.1 Introduction

The scalar wave theory that we discussed in Chapter 1 was applied to the study of light before the development of the theory of electromagnetism. At that time, it was assumed that light waves were longitudinal, in analogy with sound waves; i.e., the wave displacements were in the direction of propagation. A further assumption, that light propagated through some type of medium, was made because the scientists of that time approached all problems from a mechanistic point of view. The scalar theory was successful in explaining diffraction (see Chapter 9), but problems arose in interpretation of the effects of polarization in interference experiments (discussed in Chapter 4). Young was able to resolve the difficulties by suggesting that light waves could be transverse, like the waves on a vibrating string. Using this idea, Fresnel developed a mechanistic description of light that could explain the amount of reflected and transmitted light from the interface between two media (see Chapter 3).

Independent of this activity, the theory of electromagnetism was under development. **Michael Faraday (1791–1867)** observed in 1845 that a magnetic field would rotate the plane of polarization of light waves passing through the magnetized region. This observation led Faraday to associate light with electromagnetic radiation, but he was unable to quantify this association. Faraday attempted to develop electromagnetic theory by treating the field as lines pointing in the direction of the force that the field would exert on a test charge. The lines were given a mechanical interpretation with a tension along each line and a pressure normal to the line. **James Clerk Maxwell (1831–1879)** furnished a mathematical framework for Faraday's model in a paper read in 1864 and published a year later [1]. In this paper, Maxwell identified light as "an electromagnetic disturbance in the form of waves propagated through the electromagnetic field according to electromagnetic laws" and demonstrated that the propagation velocity of light was given by the electromagnetic properties of the material.

Maxwell was not the first to recognize the connection between the electromagnetic properties of materials and the speed of light. Kirchhoff (**Gustav Robert Kirchhoff: 1824–1887**) recognized in 1857 that the speed of light could be obtained from electromagnetic properties. Riemann (**Georg Friedrich Bernhard Riemann: 1826–1866**), in 1858, assumed that electromagnetic forces propagated at a finite velocity and derived a propagation velocity given by the electromagnetic properties of the medium. However, it was Maxwell who demonstrated that the electric and magnetic fields are waves that travel at the speed of light. It was not until 1887 that an experimental observation of electromagnetic waves, other than light, was obtained by **Heinrich Rudolf Hertz (1857–1894)**.

The classical electromagnetic theory is successful in explaining all of the experimental observations to be discussed in this book. There are, however, experiments that cannot be explained by classical wave theory, especially those conducted at short wavelengths or very low light levels. Quantum electrodynamics (QED) is capable of predicting the outcome of all optical experiments; its shortcoming is that it does not explain why or how. An excellent elementary introduction to QED has been written by Richard Feynman [2].

Modern Optics. Second Edition. B. D. Guenther.

In this chapter, we will borrow, from electromagnetic theory, Maxwell's equations and Poynting's theorem to derive properties of light waves. Details of the origins of these fundamental electromagnetic relationships are not needed for our study of light but can be obtained by consulting any electricity and magnetism text (e.g., [3]).

The basic properties that will be derived are

- the wave nature of light;
- the fact that light is a transverse wave;
- the velocity of light in terms of fundamental electromagnetic properties of materials;
- the relative magnitude of the electric and magnetic fields and relationships between the two fields;
- the energy associated with a light wave.

The concept of polarized light and a geometrical construction used to visualize its behavior will be introduced. In the seventeenth century, Hooke postulated that light waves might be transverse, but his idea was forgotten. Young and Fresnel made the same postulate in the nineteenth century and accompanied this with a theoretical description of light based on transverse waves. Forty years later, Maxwell proved that light must be a transverse wave and that **E** and **H**, for a plane wave in an isotropic medium with no free charges and no currents, are mutually perpendicular and lie in a plane normal to the direction of propagation, **k**.

Convention requires that we use the electric vector to label the direction of the electromagnetic wave's polarization. The direction of the displacement vector is called the *direction of polarization*, and the plane containing the direction of polarization and the propagation vector is called the *plane of polarization*. The selection of the electric field is not completely arbitrary—except for relativistic situations, when $v \approx c$, the interaction of the electromagnetic wave with matter will be dominated by the electric field. Both a vector and a matrix notation for describing polarization will be presented in this chapter, but details on the manipulation of light polarization will not be discussed until Chapter 14. The chapter will conclude with a discussion of the propagation of light in a conducting medium.

2.2 Maxwell's Equations

The bases of electromagnetic theory are Maxwell's equations. They allow the derivation of the properties of light. In our study of optics, we will treat these equations as axioms, but we provide the reader with a reference source here that can be consulted if information on the origin of the equations is desired.

In rationalized MKS units, Maxwell's equations are as follows.

2.2.1 Gauss's Law

2.2.1.1 Gauss's (Coulomb's) Law for the Electric Field

Coulomb's law provides a means for calculating the force between two charges (see Chapter 2 of [3]),

$$\mathbf{F} = \frac{q_0}{4\pi\varepsilon_0}\int \frac{dq}{r^2}\hat{\mathbf{n}},$$

where dq is the charge on an infinitesimal surface and $\hat{\mathbf{n}}$ is a unit vector in the direction of the line connecting the charges q_0 and dq. The electric field

$$\mathbf{E} = \frac{\mathbf{F}}{q_0}$$

is obtained using Coulomb's law (see Chapter 3 of [3]). We view this field as Michael Faraday did, as lines of flux, called lines of force, originating on positive charges and terminating on negative charges. Gauss's Law states that the quantity of charge contained within a closed surface is equal to the number of flux lines passing outward through the surface (see Chapter 4 of [3]). This view of the electric field leads to

$$\nabla \cdot \mathbf{D} = \rho, \tag{2.1}$$

where ρ is the charge density and $\mathbf{D}$ is the electric displacement (see Chapter 10, Section 5 of [3]). The use of the displacement allows the equation to be applied to any material.

2.2.1.2 Gauss's Law for the Magnetic Field

Charges at rest led to (2.1). Charges in motion, i.e., a current $\mathbf{i}$ or a current density $\mathbf{J}$, create a magnetic field $\mathbf{B}$ (see Chapter 14 of [3]). As we did for the electric field, we treat the magnetic field as flux lines, called lines of induction, and we assume that the current density is a constant so that $\nabla \cdot \mathbf{J} = 0$. This leads to (see Chapter 16 of [3])

$$\nabla \cdot \mathbf{B} = 0. \tag{2.2}$$

The zero is due to the fact that the magnetic equivalent of a single charge has never been observed.

2.2.2 Faraday's Law

The previous two equations are associated with electric and magnetic fields that are constant with respect to time. The next equation, an experimentally derived equation, deals with a magnetic field that is time-varying or equivalently a conductor moving through a static magnetic field:

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0. \tag{2.3}$$

In terms of the concept of flux, it states that an electric field around a circuit is associated with a change in the magnetic flux contained within the circuit.

2.2.3 Ampére's Law (Law of Biot and Savart)

An electric charge in motion creates a magnetic field around its path. The law of Biot and Savart allows us to calculate the magnetic field at a point located a distance $\mathbf{R}$ from a conductor carrying a current density $\mathbf{J}$. Ampère's law is the inverse relationship used to calculate the current in a conductor due to the magnetic field contained in a loop about the conductor. Neither relationship is adequate when the current is a function of time. Maxwell's major contribution to physics was to observe that the addition of a *displacement current* to Ampère's law allowed fluctuating currents to be explained. The relationship became (see Chapter 21 of [3])

$$\nabla \times \mathbf{H} = \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \tag{2.4}$$

As discussed in Appendix 2B, the constants in Maxwell's equations depend on the units used. Many optics books use c.g.s. units, which result in a form for Maxwell's equations shown in Appendix 2B.

2.2.4 Constitutive Relations

The dynamic responses of atoms and molecules in the propagation medium are taken into account through what are called the *constitutive relations*:

$$\mathbf{D} = f(\mathbf{E}),$$
$$\mathbf{J} = g(\mathbf{E}),$$
$$\mathbf{B} = h(\mathbf{H}).$$

Here we will assume that the functional relations are independent of space and time, and we will write the constitutive relations as

$$\mathbf{D} = \varepsilon\mathbf{E}, \qquad \varepsilon = \text{dielectric constant},$$
$$\mathbf{J} = \sigma\mathbf{E}, \qquad \sigma = \text{conductivity (Ohm's law)},$$
$$\mathbf{B} = \mu\mathbf{H}, \qquad \mu = \text{permeability},$$

where the constants ε, σ, and μ contain the description of the material. Later, we will explore the effects resulting from the constitutive relations having a temporal or a spatial dependence.

Often, **D** and **B** are defined as

$$\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P}, \tag{2.5}$$
$$\mathbf{B} = \mu_0\mathbf{H} + \mathbf{M},$$

where **P** is the polarization and **M** is the magnetization. This formulation emphasizes that the internal field of a material is due not only to the applied field but also due to a field created by the atoms and molecules that make up the material. We will find (2.5) useful in Chapters 7 and 15. We will not use the relationship involving **M** in this book.

By manipulating Maxwell's equations, we can obtain a number of the properties of light, such as its wave nature, the fact that it is a transverse wave, and the relationship between the **E** and **B** fields. We will make a number of simplifying assumptions about the medium in which the light is propagating to allow a quick derivation of the properties of light. Later, we will see what happens if we modify these assumptions.

2.3 Free Space

We assume that the light is propagating in a medium that we will call *free space* and that is

(1) uniform: ε and μ have the same value at all points;
(2) isotropic: ε and μ do not depend upon the direction of propagation;
(3) nonconducting: $\sigma = 0$, and thus $\mathbf{J} = 0$;
(4) free from charge: $\rho = 0$;
(5) nondispersive: ε and μ are not functions of frequency, i.e., they have no time dependence.

Our definition departs somewhat from other definitions of free space in that we include in the definition not only the vacuum, where $\varepsilon = \varepsilon_0$ and $\mu = \mu_0$, but also dielectrics, where $\sigma = 0$ but the other electromagnetic constants can have arbitrary values.

Using the above assumptions, Maxwell's equations and the constitutive relations simplify to

$$\nabla \cdot \mathbf{E} = 0, \qquad (2.6a) \qquad\qquad \nabla \cdot \mathbf{B} = 0, \qquad (2.6b)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad (2.6c) \qquad\qquad \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}, \qquad (2.6d)$$

$$\mathbf{B} = \mu \mathbf{H}, \qquad (2.6e) \qquad\qquad \varepsilon \mathbf{E} = \mathbf{D}. \qquad (2.6f)$$

These simplified equations can now be used to derive some of the basic properties of a light wave.

2.4 Wave Equation

To find how the electromagnetic wave described by (2.6) propagates in free space, Maxwell's equations must be rearranged to display explicitly the time and coordinate dependence. Using (2.6e, f), we can rewrite (2.6d) as

$$\frac{1}{\mu} \nabla \times \mathbf{B} = \varepsilon \frac{\partial \mathbf{E}}{\partial t}.$$

The curl of (2.6c) is taken and the magnetic field dependence is eliminated by using the rewritten (2.6d):

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t}\right) = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = -\frac{\partial}{\partial t}\left(\varepsilon\mu \frac{\partial \mathbf{E}}{\partial t}\right).$$

The assumption that ε and μ are independent of time allows this equation to be rewritten as

$$\nabla \times (\nabla \times \mathbf{E}) = -\varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

Using the vector identity (2A.12) from Appendix 2A, we can write

$$\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\varepsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2}.$$

Because free space is free of charge, $\nabla \cdot \mathbf{E} = 0$, giving us

$$\nabla^2 \mathbf{E} = \mu\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{2.7}$$

We can use the same procedure to obtain

$$\nabla^2 \mathbf{B} = \mu\varepsilon \frac{\partial^2 \mathbf{B}}{\partial t^2}. \tag{2.8}$$

These equations are wave equations, with the wave's velocity being given by

$$v = \frac{1}{\sqrt{\mu\varepsilon}}. \tag{2.9}$$

The connection of the velocity of light with the electric and magnetic properties of a material was one of the most important results of Maxwell's theory. In a vacuum,

$$\mu_0\varepsilon_0 = \left(4\pi \times 10^{-7}\right)\left(8.8542 \times 10^{-12}\right) = 1.113 \times 10^{-17}\ \mathrm{s^2/m^2},$$

$$\frac{1}{\sqrt{\mu_0\varepsilon_0}} = 2.998 \times 10^8\ \mathrm{m/s} = c. \tag{2.10}$$

In a material, the velocity of light is less than c. We can characterize a material by defining the *index of refraction*, the ratio of the speed of light in a vacuum to its speed in a medium:

$$n = \frac{c}{v} = \sqrt{\frac{\varepsilon\mu}{\varepsilon_0\mu_0}}. \tag{2.11}$$

The data in Table 2.1 demonstrate that if magnetic materials are not considered, then $\mu/\mu_0 \approx 1$, so that

$$n = \sqrt{\frac{\varepsilon}{\varepsilon_0}}\ .$$

Table 2.1 *Representative magnetic permeabilities*

Material	μ/μ_0	Class
Silver	0.99998	Diamagnetic
Copper	0.99999	Diamagnetic
Water	0.99999	Diamagnetic
Air	1.00000036	Paramagnetic
Aluminum	1.000021	Paramagnetic
Iron	5000	Ferromagnetic
Nickel	600	Ferromagnetic

Table 2.2 *Selected indices of refraction*

Material	n (yellow light)	$\sqrt{\varepsilon/\varepsilon_0}$ (static)
Air	1.000294	1.000295
CO_2	1.000449	1.000473
C_6H_6 (benzene)	1.482	1.489
He	1.000036	1.000034
H_2	1.000131	1.000132

The data displayed in Table 2.2 demonstrate that, at least for some materials, the theory agrees with experimental results. The materials whose indices are listed in Table 2.2 have been specially selected to demonstrate good agreement; we will see in Chapter 7 that the assumption that ε, μ, and σ are independent of frequency results in a theory that neglects the response time of the system to the electromagnetic signal.

2.5 Transverse Waves

Hooke postulated, in the seventeenth century, that light waves might be transverse, but his idea was forgotten. In the nineteenth century, Young and Fresnel made the same postulate and provided a theoretical description of light based on transverse waves. Forty years later, Maxwell proved that light must be a transverse wave. We can demonstrate the transverse nature of light by substitution of the plane wave solution of the wave equation into Gauss's law:

$$\nabla \cdot \mathbf{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = 0.$$

To complete the demonstration, we consider the divergence of the electric component of the plane wave. We will examine only the x-coordinate of the divergence in detail:

$$\begin{aligned}\frac{\partial E_x}{\partial x} &= \frac{\partial}{\partial x}\left(E_{0x}e^{i(\omega t - \mathbf{k}\cdot\mathbf{r} + \phi)}\right) = iE_{0x}e^{i(\omega t - \mathbf{k}\cdot\mathbf{r} + \phi)}\frac{\partial}{\partial x}(\omega t - k_x x - k_y y - k_z z + \phi)\\ &= -ik_x E_x.\end{aligned}$$

We easily obtain similar results for E_y and E_z, allowing the divergence of $\mathbf{E}$ to be rewritten as a dot product of $\mathbf{k}$ and $\mathbf{E}$. Gauss's law for the electric field states that the divergence of $\mathbf{E}$ is zero, which for a plane wave can be written

$$\nabla \cdot \mathbf{E} = -i\mathbf{k} \cdot \mathbf{E} = 0. \tag{2.12}$$

If the dot product of two vectors $\mathbf{E}$ and $\mathbf{k}$, is zero, then the vectors $\mathbf{E}$ and $\mathbf{k}$ must be perpendicular [see (2A.1) in Appendix 2A]. In the same manner, substituting the plane wave into $\nabla \cdot \mathbf{B} = 0$ yields $\mathbf{k} \cdot \mathbf{B} = 0$. Therefore, Maxwell's equations require light to be a transverse wave; i.e., the vector displacements $\mathbf{E}$ and $\mathbf{B}$ are perpendicular to the direction of propagation, $\mathbf{k}$.

2.6 Interdependence of E and B

The electric and magnetic fields are not independent, as we can see by continuing our examination of the plane wave solutions of Maxwell's equations. First, let us calculate several derivatives of the plane wave. We will need

$$\begin{aligned}\frac{\partial \mathbf{B}}{\partial t} &= \frac{\partial}{\partial t}\mathbf{B}_0 e^{i(\omega t - \mathbf{k}\cdot\mathbf{r}+\phi)} = i\mathbf{B}\frac{\partial}{\partial t}(\omega t - \mathbf{k}\cdot\mathbf{r}+\phi) \\ &= i\omega\mathbf{B}\end{aligned} \tag{2.13}$$

and, similarly,

$$\frac{\partial \mathbf{E}}{\partial t} = i\omega\mathbf{E}. \tag{2.14}$$

A simple expression for the curl of $\mathbf{E}$, $\nabla \times \mathbf{E}$, can be obtained when we use the derivatives just calculated. The expression for the curl of $\mathbf{E}$ is given by (2A.7) from Appendix 2A and is rewritten here:

$$\nabla \times \mathbf{E} = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)\hat{\mathbf{k}}.$$

The terms making up the x-component of the curl are

$$\frac{\partial E_z}{\partial y} = E_{0z}\frac{\partial}{\partial y}e^{i(\omega t - \mathbf{k}\cdot\mathbf{r}+\phi)} = -ik_y E_z$$

and

$$\frac{\partial E_y}{\partial z} = -ik_z E_y.$$

By evaluating each component, we find that the curl of $\mathbf{E}$ for a plane wave is

$$\nabla \times \mathbf{E} = -i\mathbf{k} \times \mathbf{E}. \tag{2.15}$$

A similar derivation leads to the curl of $\mathbf{B}$ for a plane wave:

$$\nabla \times \mathbf{B} = -i\mathbf{k} \times \mathbf{B}. \tag{2.16}$$

With these vector operations on a plane wave defined, we can evaluate (2.6c) for a plane wave. The left-hand side of

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

is replaced with (2.15) and the right-hand side by (2.13), resulting in an equation connecting the electric and magnetic fields:

$$-i\mathbf{k} \times \mathbf{E} = -i\omega\mathbf{B}.$$

Using the relationship between ω and **k** given by (1.2) from Chapter 1 and the relationship for the wave velocity in terms of the electromagnetic properties of the material, (2.9), we can write

$$\frac{\sqrt{\mu\varepsilon}}{k}\mathbf{k} \times \mathbf{E} = \mathbf{B}. \tag{2.17}$$

A second relationship between the magnetic and electric fields can be generated by using the same procedure to rewrite

$$\nabla \times \mathbf{B} = \mu\varepsilon \frac{\partial \mathbf{E}}{\partial t}$$

for a plane wave as

$$-i\mathbf{k} \times \mathbf{B} = i\varepsilon\mu\omega\mathbf{E};$$

that is,

$$\frac{1}{k\sqrt{\mu\varepsilon}}\mathbf{k} \times \mathbf{B} = -\mathbf{E}. \tag{2.18}$$

From the definition of the cross product given by (2A.2) in Appendix 2A, we see that the electric and magnetic fields are perpendicular to each other, are in phase, and form a right-handed coordinate system with propagation direction **k** (see Figure 2.1).

If we are only interested in the magnitude of the two fields, we can use (2.11) to write

$$n|\mathbf{E}| = c|\mathbf{B}|. \tag{2.19}$$

In a vacuum, we take $n = 1$ in (2.19). For our plane wave, the ratio of the field magnitudes is

$$\frac{|\mathbf{E}|}{|\mathbf{H}|} = \sqrt{\frac{\mu}{\varepsilon}}. \tag{2.20}$$

This ratio has units of ohms ($\mu \Rightarrow \mathrm{ml/Q^2}$, $\varepsilon \Rightarrow \mathrm{Q^2t^2/ml^3}$, and $\Omega \Rightarrow \mathrm{ml^2/Q^2t}$) and is called the impedance of the medium. In a vacuum,

$$Z_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}} = 377\ \Omega.$$

When the ratio is a real quantity, as it is here, **E** and **H** are in phase.

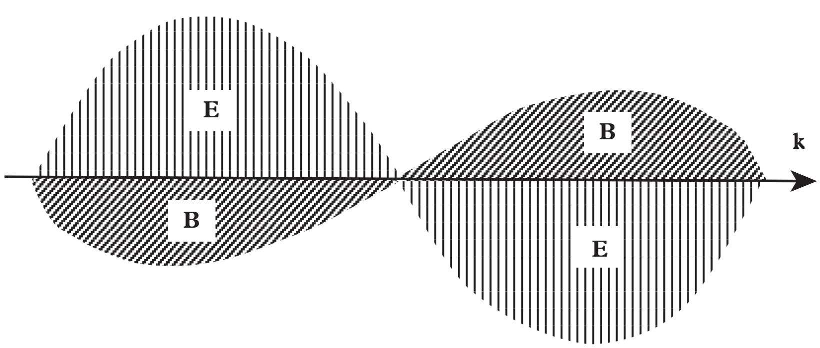

Figure 2.1 *Graphical representation of an electromagnetic plane wave. Note that* **E** *and* **B** *are perpendicular to each other and individually perpendicular to the propagation vector* **k**, *are in phase, and form a right-handed coordinate system as required by (2.17) and (2.18).*

2.7 Energy Density and Flow

We saw in our discussion of waves propagating along strings that the power transmitted by a wave is proportional to the square of the amplitude of the wave. Any text on electromagnetic theory (see, e.g., Chapter 21 of [3]) demonstrates that the energy density (in J/m^3) associated with an electromagnetic wave is given by

$$U = \frac{\mathbf{D} \cdot \mathbf{E} + \mathbf{B} \cdot \mathbf{H}}{2}. \tag{2.21}$$

We can simplify (2.21) by using the simple constitutive relations $\mathbf{D} = \varepsilon \mathbf{E}$ and $\mathbf{B} = \mu \mathbf{H}$, if they apply to the propagation medium:

$$U = \frac{1}{2}\left(\varepsilon E^2 + \frac{B^2}{\mu}\right) = \frac{1}{2}\left(\varepsilon + \frac{1}{\mu c^2}\right) E^2.$$

In a vacuum, further simplification is possible:

$$U = \varepsilon_0 E^2 = \frac{B^2}{\mu_0}.$$

John Henry Poynting (1852–1914), an English professor of physics at Mason Science College, now the University of Birmingham, demonstrated that the presence of both an electric and a magnetic field at the same point in space results in a flow of the field energy. This fact is called the Poynting theorem, and the flow is completely described by the *Poynting vector*

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}. \tag{2.22}$$

The Poynting vector has units of J/(m^2·s). We will use a plane wave to determine some of the properties of this vector. Since $\mathbf{S}$ will involve terms quadratic in $\mathbf{E}$, it will be necessary to use the real form of $\mathbf{E}$ (see Problem 1.4). We have

$$\mathbf{H} = \frac{\mathbf{B}}{\mu} = \frac{\sqrt{\mu\varepsilon}}{\mu k}\mathbf{k} \times \mathbf{E},$$

where

$$\mathbf{E} = \mathbf{E}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi).$$

Then,

$$\begin{aligned}\mathbf{S} &= \frac{\sqrt{\mu\varepsilon}}{\mu k}\mathbf{E}_0 \times (\mathbf{k} \times \mathbf{E}_0) \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi) \\ &= \frac{n}{\mu c} |\mathbf{E}_0|^2 \frac{\mathbf{k}}{k} \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi).\end{aligned} \tag{2.23}$$

Note that the energy is flowing in the direction of propagation (indicated by the unit vector $\mathbf{k}/k$).

We normally do not detect **S** at the very high frequencies associated with light ($\approx 10^{15}$ Hz) but rather detect a temporal average of **S** taken over a time T determined by the response time of the detector used. We must obtain the time average of **S** to relate theory to actual measurements. The time average of **S** is called the *flux density* and has units of W/m^2. We will call this quantity the ***intensity*** of the light wave,

$$I = |\langle \mathbf{S} \rangle| = \left| \frac{1}{T} \int_{t_0}^{t_0+T} \mathbf{A} \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi)\, dt \right|, \tag{2.24}$$

where we have defined

$$\mathbf{A} = \frac{n}{\mu c} |\mathbf{E}_0|^2 \frac{\mathbf{k}}{k}$$

to simplify the notation.

The units used for the flux density are a confusing mess in optics. One area of optics is interested in measuring the physical effects of light, and the measurement of energy is called *radiometry*. In radiometry, the flux density is called the *irradiance*, with units of W/m^2. Another area of optics is interested in the psychophysical effects of light, and the measurement of energy is called *photometry*. For this group, the flux density is called *illuminance*, with units of a lumen/m^2 (lm/m^2) or a lux. Each of these two group has its own set of units, but both desire to measure the energy flow of a field that is not well defined in frequency or phase. Much research in modern optics belong to a third area that is associated with the use of a light source that has both a well-defined frequency and a well-defined phase—the laser. In this area of optics, common usage defines the time-averaged flux density as the intensity. In this book, all of the waves discussed are uniquely defined in terms of the electric field and the electromagnetic properties of the material in which the wave is propagating. To emphasize that the results of our theory are immediately applicable only to a light source with a well-defined frequency and phase, we will use the term *intensity* for the magnitude of the Poynting vector.

We will assume that **k** is independent of time over the period T:

$$\langle \mathbf{S} \rangle = \frac{\mathbf{A}}{\omega T} \int_{t_0\omega}^{(t_0+T)\omega} \cos^2(\omega t - \mathbf{k} \cdot \mathbf{r} + \phi)\, d(\omega t).$$

Using the trigonometric identity

$$\cos^2\theta = \frac{1}{2}(1 + \cos 2\theta)$$

and evaluating the integral results in the expression

$$\langle \mathbf{S} \rangle = \frac{\mathbf{A}}{2} + \frac{\mathbf{A}}{4\omega T}[\sin 2(\omega t_0 + \omega T - \mathbf{k} \cdot \mathbf{r} + \phi) - \sin 2(\omega t_0 - \mathbf{k} \cdot \mathbf{r} + \phi)]. \tag{2.25}$$

The largest value that the term in square brackets can assume is 2. The period T is the response time of the detector to the light wave. Normally, it is much longer than the period of light oscillations, so $\omega T \gg 1$ and we can neglect the second term in (2.25). As an example, suppose our detection system has a 1 GHz bandwidth, yielding a response time of $T = 10^{-9}$ s (the reciprocal of the bandwidth). Green light has a frequency of $\nu = 6 \times 10^{14}$ Hz or $\omega \approx 4 \times 10^{15}$ rad/s. With these values, $\omega T = 4 \times 10^5$, and the neglected term would be no larger than 10^{-6} of the first term. Therefore, in optics, the assumption that $\omega T \gg 1$ is reasonable and allows the average Poynting vector to be written as

$$\langle \mathbf{S} \rangle = \frac{\mathbf{A}}{2} = \frac{n}{2\mu c} |\mathbf{E}_0|^2 \frac{\mathbf{k}}{k}. \tag{2.26}$$

The energy per unit time per unit area depends on the square of the amplitude of the wave.

The energy calculation was done with plane waves of **E** and **H** that are in phase. We will later see that materials with nonzero conductivity $\sigma \neq 0$ will yield a complex impedance because **E** and **H** are no longer in phase. If the two waves are 90° out of phase, then the integral in (2.24) will contain $\sin x \ \cos x$ as its integrand, resulting in $\langle \mathbf{S} \rangle = 0$. Therefore, no energy is transmitted.

In quantum mechanics, the energy of light is carried by discrete particles called photons. If the light has a frequency ν, then the energy of a photon is $h\nu$. The intensity of the light is equal to the number of photons, striking unit area in unit time, N, multiplied by the energy of an individual photon:

$$I = Nh\nu.$$

The intensity of a 10 mW HeNe laser beam, 2 mm in diameter, is

$$I = \frac{\text{power}}{\text{area}} = \frac{10^{-2}}{\pi \left(10^{-3}\right)^2} = 3.18 \times 10^3 \text{ W m}^{-2}.$$

The number of photons in this beam can be calculated once we know that the wavelength of the light is 632.8 nm:

$$N = \frac{I}{h\nu} \cdot \text{area} = \frac{10^{-2}\lambda}{hc} = \frac{10^{-2}\left(632.8 \times 10^{-9}\right)}{\left(6.6 \times 10^{-34}\right)\left(3 \times 10^8\right)} = 32 \times 10^{15}.$$

We can get a perspective on how large this number is by comparing it with the number of molecules in a mole of a molecule, i.e., Avogadro's constant

$$N_A = 6.02214129 \times 10^{23}.$$

For carbon-12, a mole is 12 g.

The energy crossing a unit area A in a time Δt is contained in a volume $A(v\Delta t)$ (in a vacuum $v = c$), as shown in Figure 2.2. To find the magnitude of this energy, we must multiply this volume by the average energy density $|\langle \mathbf{S} \rangle|$. Thus, we expect the energy flow to be given by

$$|\langle \mathbf{S} \rangle| = \frac{\text{energy}}{A\Delta t} \propto \frac{Av\Delta t \langle U \rangle}{A\Delta t} = v \langle U \rangle.$$

We may use the definitions of the wave velocity

$$v = \frac{1}{\sqrt{\mu\varepsilon}}$$

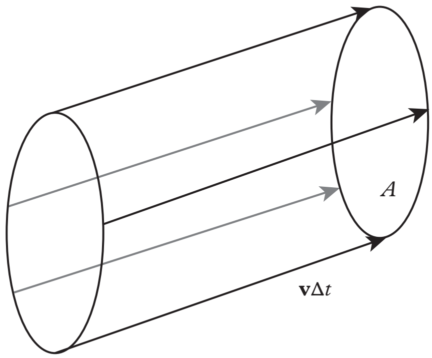

Figure 2.2 *The energy of a wave crossing a unit area A in a time Δt.*

and index of refraction $n = c/v$ to rewrite (2.26) as

$$|\langle \mathbf{S} \rangle| = \frac{\varepsilon v E_0^2}{2} = v\,\langle U \rangle, \tag{2.27}$$

giving the expected result that the energy is flowing through space at the speed of light in the medium. The relationship defined by (2.27), (energy flow) = (wave velocity) · (energy density), is a general property of waves.

At the Earth's surface, the flux density of full sunlight is 1.34×10^3 J/(m^2·s). It is not completely correct to do so, but if we associate this flux with the time average of the Poynting vector, then the electric field associated with the sunlight is $E_0 = 10^3$ V/m.

Our discussion of the time average of the Poynting vector provides an opportunity to discover one of the advantages of the use of complex notation. To obtain the time average of the product of two waves A and B, where

$$A = \mathcal{R}e\{\tilde{A}\} = \mathcal{R}e\{A_0 e^{i(\omega t + \phi_1)}\},$$
$$B = \mathcal{R}e\{\tilde{B}\} = \mathcal{R}e\{B_0 e^{i(\omega t + \phi_2)}\},$$

we use

$$\mathcal{R}e\{\tilde{z}\} = x = r\cos\phi = \frac{\tilde{z} + \tilde{z}^*}{2},$$
$$\mathcal{I}m\{\tilde{z}\} = y = r\sin\phi = \frac{\tilde{z} - \tilde{z}^*}{2i}$$

to write the average over one period as

$$\langle AB \rangle = \frac{1}{T}\int_0^T \left(\frac{\tilde{A} + \tilde{A}^*}{2}\right)\left(\frac{\tilde{B} + \tilde{B}^*}{2}\right) dt,$$
$$(\tilde{A} + \tilde{A}^*)(\tilde{B} + \tilde{B}^*) = \tilde{A}\tilde{B} + \tilde{A}^*\tilde{B}^* + \tilde{A}\tilde{B}^* + \tilde{A}^*\tilde{B},$$

where

$$\tilde{A}\tilde{B} = A_0 B_0 e^{i(2\omega t+\phi_1+\phi_2)},$$

$$\tilde{A}^*\tilde{B}^* = A_0 B_0 e^{-i(2\omega t+\phi_1+\phi_2)}.$$

The time averages of the latter two terms are zero, and we are left with

$$\langle AB\rangle = \frac{1}{T}\int_0^T \frac{\tilde{A}\tilde{B}^* + \tilde{A}^*\tilde{B}}{4}\,dt.$$

Again using $\mathcal{Re}\{\tilde{z}\} = (\tilde{z} + \tilde{z}^*)/2$, we may rewrite this as

$$\langle AB\rangle = \frac{1}{2}\mathcal{Re}\left\{\tilde{A}\tilde{B}^*\right\}. \tag{2.28}$$

The reader may find this quite general relation easier to use than performing an integration such as (1.24).

2.8 Polarization

The displacement of a transverse wave is a vector quantity. We must therefore specify the frequency, phase, and direction of the wave along with the magnitude and direction of the displacement. The direction of the displacement vector is called the *direction of polarization*, and the plane containing the direction of polarization and the propagation vector is called the *plane of polarization*. This quantity has the same name as the field quantity introduced in (2.5). Because the two terms describe completely different physical phenomena, there should be no danger of confusion.

From our study of Maxwell's equations, we know that **E** and **H**, for a plane wave in free space, are mutually perpendicular and lie in a plane normal to the direction of propagation, **k**. We also know that, given one of the two vectors, we can use (2.17) to obtain the other. Convention requires that we use the electric vector to label the direction of the electromagnetic wave's polarization. The selection of the electric field is not completely arbitrary.

The electric field of the electromagnetic wave acts on a charged particle in the material with a force

$$\mathbf{F}_E = q\mathbf{E}. \tag{2.29}$$

This force accelerates the charged particle to a velocity v in a direction transverse to the direction of light propagation and parallel to the electric field. The moving charge interacts with the magnetic field of the electromagnetic wave with a force

$$\mathbf{F}_H = q(\mathbf{v}\times\mathbf{B}), \tag{2.30}$$

parallel to the propagation vector. We can write the ratio of the forces on a moving charge in an electromagnetic field due to the electric and magnetic fields as

$$\frac{F_E}{F_H} = \frac{eE}{evB}.$$

We can replace B using (2.19) to obtain

$$\frac{F_E}{F_H} = \frac{c}{nv}, \tag{2.31}$$

where v is the velocity of the moving charge. Assuming that a charged particle is traveling in air at the speed of sound, so that $v = 335$ m/s, the force due to the electric field of a light wave on that particle will therefore be 8.9×10^5 times larger than the force due to the magnetic field. The size of this numbers demonstrates that, except in relativistic situations, when $v \approx c$, the interaction of an electromagnetic wave with matter will be dominated by the electric field.

A conventional vector notation is used to describe the polarization of a light wave; however, to visualize the behavior of the electric field vector as light propagates, a geometrical construction is useful. The geometrical construction, called a Lissajous figure, describes the path followed by the tip of the electric field vector.

2.8.1 Polarization Ellipse

Assume that a plane wave is propagating in the z-direction and that the electric field, determining the direction of polarization, is oriented in the (x, y) plane. In complex notation, the plane wave is given by

$$\tilde{\mathbf{E}} = \mathbf{E}_0 e^{i(\omega t - \mathbf{k}\cdot\mathbf{r} + \phi)} = \mathbf{E}_0 e^{i(\omega t - kz + \phi)}.$$

This wave can be written in terms of the x- and y-components of $\mathbf{E}_0$:

$$\tilde{\mathbf{E}} = E_{0x} e^{i(\omega t - kz + \phi_1)} \hat{\mathbf{i}} + E_{0y} e^{i(\omega t - kz + \phi_2)} \hat{\mathbf{j}}. \tag{2.32}$$

(To prevent errors, we will use only the real part of $\mathbf{E}$ for manipulation.) We divide each component of the electric field by its maximum value so that the problem is reduced to one of the following two sinusoidally varying unit vectors:

$$\frac{E_x}{E_{0x}} = \cos(\omega t - kz + \phi_1) = \cos(\omega t - kz)\cos\phi_1 - \sin(\omega t - kz)\sin\phi_1,$$
$$\frac{E_y}{E_{0y}} = \cos(\omega t - kz)\cos\phi_2 - \sin(\omega t - kz)\sin\phi_2.$$

When these unit vectors are added together, the result will be a set of figures called *Lissajous figures* (**Jules Antoine Lissajous: 1822–1880**). The geometrical construction shown in Figure 2.3 can be used to visualize the generation of the Lissajous figure. The harmonic motion along the x-axis is found by projecting a vector rotating around a circle of diameter E_{0x} onto the x-axis. The harmonic motion, along the y-axis is generated the same way using a circle of diameter E_{0y}. The resulting x- and y-components are added to obtain $\mathbf{E}$. In Figure 2.3, the two harmonic oscillators both have the same frequency, $\omega t - kz$, but differ in phase by

$$\delta = \phi_2 - \phi_1 = -\frac{\pi}{2}.$$

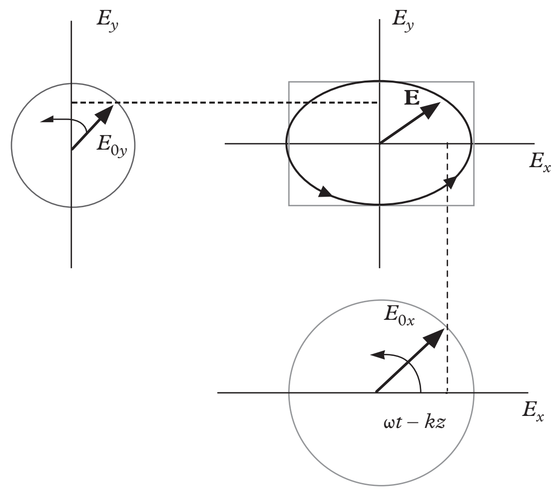

Figure 2.3 *Geometrical construction showing how the Lissajous figure is constructed from harmonic motion along the x- and y-coordinate axes. The harmonic motion along each coordinate axis is created by projecting a vector rotating around a circle onto the axis.*

The tip of the electric field $\mathbf{E}$ in Figure 2.3 traces out an ellipse, with its axes aligned with the coordinate axes. To determine the direction of the rotation of the vector, assume that $\phi_1 = 0$, $\phi_2 = -\pi/2$, and $z = 0$, so that

$$\frac{E_x}{E_{0x}} = \cos\omega t, \quad \frac{E_y}{E_{0y}} = \sin\omega t$$

$$\mathbf{E} = \left(\frac{E_x}{E_{0x}}\right)\hat{\mathbf{i}} + \left(\frac{E_y}{E_{0y}}\right)\hat{\mathbf{j}}.$$

Table 2.3 *Rotating electric field vector* $\mathbf{E}$

ωt	$\mathbf{E}$
0	$\hat{\mathbf{i}}$
$\frac{\pi}{4}$	$\frac{1}{\sqrt{2}}\left(\hat{\mathbf{i}}+\hat{\mathbf{j}}\right)$
$\frac{\pi}{2}$	$\hat{\mathbf{j}}$
$\frac{3\pi}{4}$	$\frac{1}{\sqrt{2}}\left(-\hat{\mathbf{i}}+\hat{\mathbf{j}}\right)$
π	$-\hat{\mathbf{i}}$

The normalized vector $\mathbf{E}$ can easily be evaluated at a number of values of ωt to discover the direction of rotation. Table 2.3 shows the value of the vector as ωt increases.

The rotation of the vector $\mathbf{E}$ in Figure 2.3 is seen to be in a counterclockwise direction, moving from the positive x-direction, to the y-direction, and finally to the negative x-direction.

To obtain the equation for the Lissajous figure, we eliminate the dependence of the unit vectors on $\omega t - kz$. First, we multiply the two equations by $\sin\phi_2$ and $\sin\phi_1$, respectively, and then subtract the resulting equations. Second, we multiply the two equations by $\cos\phi_2$ and $\cos\phi_1$, respectively, and then subtract the resulting equations. These two operations yield the following pair of equations;

$$\frac{E_x}{E_{0x}}\sin\phi_2 - \frac{E_y}{E_{0y}}\sin\phi_1 = \cos(\omega t - kz)\,[\cos\phi_1\sin\phi_2 - \sin\phi_1\cos\phi_2],$$

$$\frac{E_x}{E_{0x}}\cos\phi_2 - \frac{E_y}{E_{0y}}\cos\phi_1 = \sin(\omega t - kz)\,[\cos\phi_1\sin\phi_2 - \sin\phi_1\cos\phi_2].$$

The term in square brackets can be simplified using the trigonometric identity

$$\sin\delta = \sin(\phi_2 - \phi_1) = \cos\phi_1\sin\phi_2 - \sin\phi_1\cos\phi_2.$$

After replacing the term in square brackets by $\sin\delta$, the two equations are squared and added, yielding the equation for the Lissajous figure:

$$\left(\frac{E_x}{E_{0x}}\right)^2 + \left(\frac{E_y}{E_{0y}}\right)^2 - \left(\frac{2E_xE_y}{E_{0x}E_{0y}}\right)\cos\delta = \sin^2\delta. \tag{2.33}$$

The trigonometric identity

$$\cos\delta = \cos(\phi_2 - \phi_1) = \cos\phi_1\cos\phi_2 + \sin\phi_1\sin\phi_2$$

was also used here to further simplify (2.33).

Equation (2.33) has the same form as the equation of a conic,

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

The conic here can be seen to be an ellipse because, from (2.33),

$$B^2 - 4AC = \frac{4}{E_{0x}^2E_{0y}^2}(\cos^2\delta - 1) < 0.$$

This ellipse is called the *polarization ellipse*. Its orientation with respect to the x-axis is given by

$$\tan 2\theta = \frac{B}{A - C} = \frac{2E_{0x}E_{0y}\cos\delta}{E_{0x}^2 - E_{0y}^2}. \tag{2.34}$$

If $A = C$ and $B \neq 0$ then $\theta = 45°$. When $\delta = \pm\pi/2$, then $\theta = 0°$, as shown in Figure 2.3.

The tip of the resultant electric field vector obtained from (2.34) traces out the polarization ellipse in the plane normal to **k**, as predicted by (2.33). A generalized polarization ellipse is shown in Figure 2.4. The x- and y-coordinates of the electric field are bounded by $\pm E_{0x}$ and $\pm E_{0y}$. The rectangle in Figure 2.4 illustrates these limits. The component of the electric field along the major axis of the ellipse is

$$E_M = E_x\cos\theta + E_y\sin\theta$$

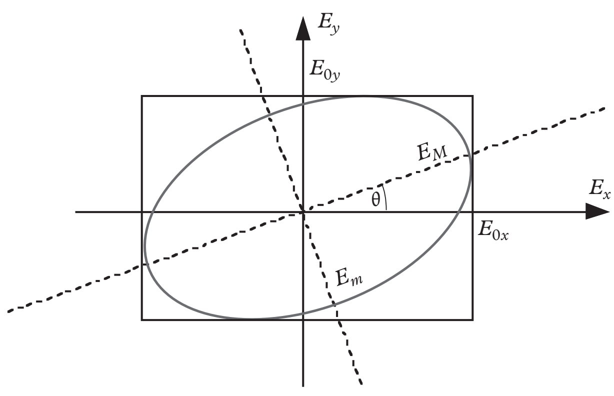

Figure 2.4 *General form of the ellipse described by (2.33).*

and that along the minor axis is

$$E_m = -E_x \sin\theta + E_y \cos\theta,$$

where θ is obtained from (2.34). The ratio of the length of the minor axis to that of the major axis is related to the ellipticity φ, which measures the amount of deviation of the ellipse from a circle:

$$\tan\varphi = \pm\left(\frac{E_m}{E_M}\right) = \frac{E_{0x}\sin\phi_1 \sin\theta - E_{0y}\sin\phi_2\cos\theta}{E_{0x}\cos\phi_1\cos\theta + E_{0y}\cos\phi_2\sin\theta}. \tag{2.35}$$

To find the time dependence of the vector $\mathbf{E}$, we rewrite (2.32) in complex form:

$$\tilde{\mathbf{E}} = e^{i(\omega t - kz)}\left(\hat{\mathbf{i}}E_{0x}e^{i\phi_1} + \hat{\mathbf{j}}E_{0y}e^{i\phi_2}\right). \tag{2.36}$$

This equation shows explicitly that the electric vector moves about the ellipse in a sinusoidal motion.

By specifying the parameters that characterize the polarization ellipse (θ and φ), we completely characterize the polarization of a wave. A review of two special cases will aid in understanding the polarization ellipse.

2.8.1.1 Linear Polarization

First consider the cases when $\delta = 0$ or π. Then (2.33) becomes

$$\left(\frac{E_x}{E_{0x}}\right)^2 + \left(\frac{E_y}{E_{0y}}\right)^2 \mp \frac{2E_xE_y}{E_{0x}E_{0y}} = 0.$$

The ellipse collapses into a straight line with slope E_{0y}/E_{0x}. The equation of the straight line is

$$\frac{E_x}{E_{0x}} = \mp\frac{E_y}{E_{0y}}.$$

Figure 2.5 displays the straight-line Lissajous figures for the two phase differences. The θ-parameter of the ellipse is the slope of the straight line,

$$\tan\theta = \frac{E_{0y}}{E_{0x}},$$

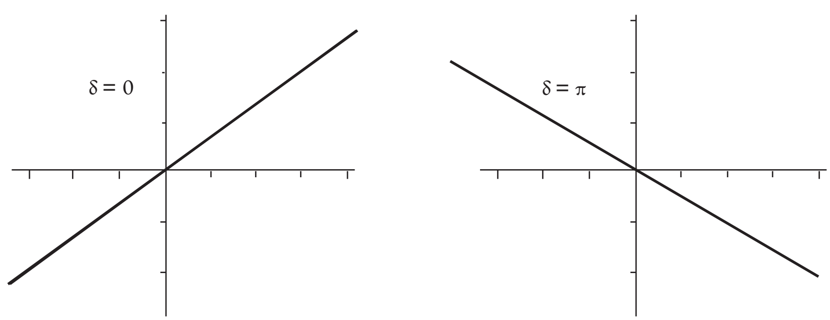

Figure 2.5 *Lissajous figures for phase differences between the y- and x-components of oscillation of 0 and π.*

resulting in the value of (2.34) being given by

$$\tan 2\theta = \frac{2\tan\theta}{1-\tan^2\theta} = \frac{2E_{0x}E_{0y}}{E_{0x}^2 - E_{0y}^2}.$$

The φ-parameter is given by (2.35) as $\tan\varphi = 0$.

The time dependence of the $\mathbf{E}$ vector shown in Figure 2.5 is given by (2.36). The real component is

$$\mathbf{E} = \left(E_{0x}\hat{\mathbf{i}} \pm E_{0y}\hat{\mathbf{j}}\right)\cos(\omega t - kz).$$

At a fixed point in space, the x- and y-components oscillate in phase (or 180° out of phase) according to the equation

$$\mathbf{E} = \left(E_{0x}\hat{\mathbf{i}} \pm E_{0y}\hat{\mathbf{j}}\right)\cos(\omega t - \phi).$$

The electric vector undergoes simple harmonic motion along the line defined by E_{0x} and E_{0y}. At a fixed time, the electric field varies sinusoidally along the propagation path (the z-axis) according to the equation

$$\mathbf{E} = \left(E_{0x}\hat{\mathbf{i}} \pm E_{0y}\hat{\mathbf{j}}\right)\cos(\phi - kz).$$

This light is said to be *linearly polarized.*

2.8.1.2 Circular Polarization

The second case occurs when $E_{0x} = E_{0y} = E_0$ and $\delta = \pm\pi/2$. From (2.33),

$$\left(\frac{E_x}{E_0}\right)^2 + \left(\frac{E_y}{E_0}\right)^2 = 1.$$

The ellipse becomes a circle as shown in Figure 2.6. For this polarization, $\tan 2\theta$ is indeterminate and $\tan\varphi = 1$.

From (2.38), the temporal behavior is given by

$$\mathbf{E} = E_0[\cos(\omega t - kz)\,\hat{\mathbf{i}} \pm \sin(\omega t - kz)\,\hat{\mathbf{j}}].$$

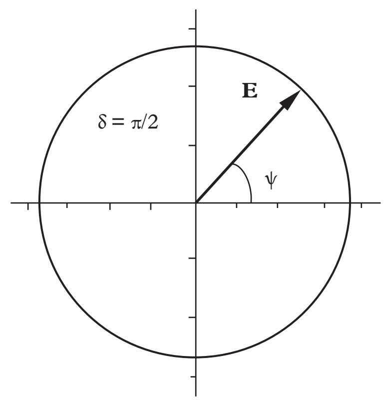

Figure 2.6 *Lissajous figure for the case when the phase differences between the y- and x-components of oscillation differ by $\pm\pi/2$ and the amplitudes of the two components are equal. The tip of the electric field vector moves along the circular path shown in the figure.*

The time dependence of the angle ψ that the $\mathbf{E}$ field makes with the x-axis in Figure 2.6 can be obtained by finding the tangent of ψ:

$$\tan\psi = \frac{E_y}{E_x} = \pm\frac{\sin(\omega t - kz)}{\cos(\omega t - kz)} = \pm\tan(\omega t - kz).$$

The interpretation of this result is that at a fixed point in space, the $\mathbf{E}$ vector rotates in a clockwise direction if $\delta = \pi/2$ and a counterclockwise direction if $\delta = -\pi/2$.

In particle physics, the light would be said to have a negative helicity if it rotated in a clockwise direction. If we look at the source, the electric vector seems to follow the threads of a left-handed screw, agreeing with the nomenclature that left-handed quantities are negative.

However, in optics, the light that rotates clockwise as we view it traveling toward us from the source is said to be *right-circularly polarized.* The counterclockwise-rotating light is *left-circularly polarized.*

The association of right-circularly polarized light with "right-handedness" in optics came about by looking at the path of the electric vector in space at a fixed time: then $\tan\psi = \tan(\phi - kz)$; see Figure 2.7. As shown in Figure 2.7, right-circularly polarized light at a fixed time seems to spiral in a counterclockwise fashion along the z-direction, following the threads of a right-handed screw.

This motion can be generalized to include elliptically polarized light when $E_{0x} \neq E_{0y}$. Figure 2.3 schematically displays the generation of the Lissajous figure for the case $\delta = \pi/2$ but with unequal values of E_{0x} and E_{0y}. Figure 2.8 shows two calculated Lissajous figures. If the electric vector moves around the ellipse in a clockwise direction, as we face the source, then the phase difference and ellipticity are

$$0 \leq \delta \leq \pi \qquad \text{and} \qquad 0 < \varphi < \frac{\pi}{4},$$

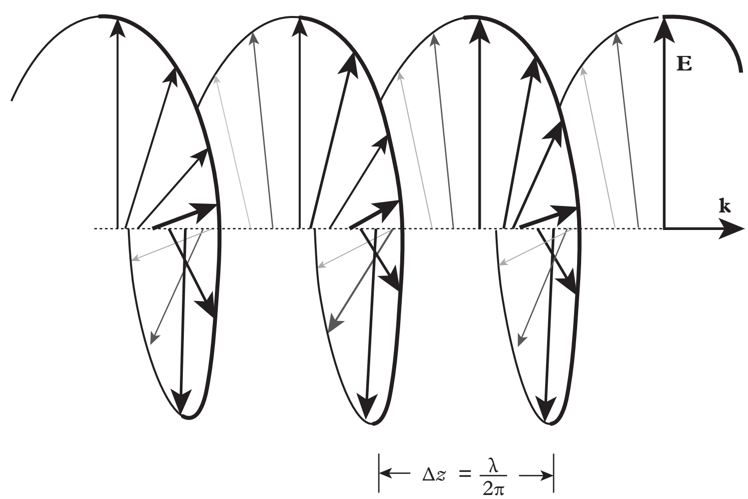

Figure 2.7 *The path of the electric vector of right-circularly polarized light at a fixed time.*

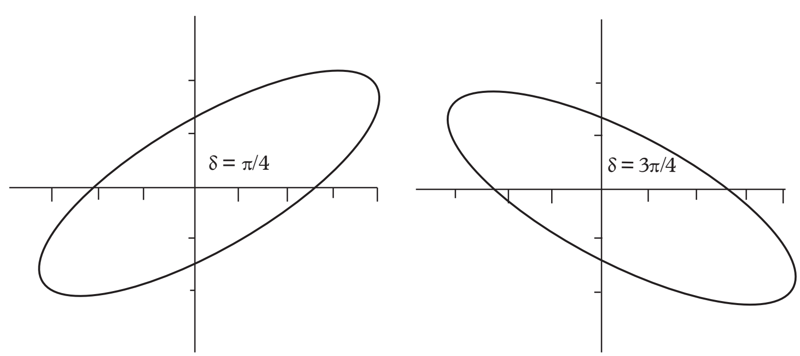

Figure 2.8 *Lissajous figures for elliptically polarized light. These were calculated with* $E_{0x} = 0.75$ *and* $E_{0y} = 0.25$.

and the polarization is right-handed. If the motion of the electric vector is moving in a counterclockwise direction, then the phase difference and ellipticity are

$$-\pi \leq \delta \leq 0 \qquad \text{and} \qquad -\frac{\pi}{4} < \varphi < 0.$$

The orientation of either ellipse, with respect to the x-axis, is given by (2.34) and depends upon the relative magnitudes of E_{0x} and E_{0y}.

The procedure used to decompose an arbitrary polarization into polarizations parallel to two axes of a Cartesian coordinate system is a technique used extensively in vector algebra to simplify mathematical calculations. According to the mathematical formalism associated with this technique, the polarization is described in terms of a set of basis vectors $\mathbf{e}_i$. An arbitrary polarization would be expressed as

$$\mathbf{E} = \sum_{i=1}^{2} a_i \mathbf{e}_i. \tag{2.37}$$

The set of basis vectors $\mathbf{e}_i$ are orthonormal, i.e.,

$$\mathbf{e}_i \mathbf{e}_j^* = \delta_{ij} = \begin{cases} 1 & (i = j), \\ 0 & (i \neq j), \end{cases}$$

where we have assumed that the basis vectors could be complex. We mention this mathematical formalism because an identical formalism is encountered in elementary particle physics, where it is used to describe spin [4].

In a Cartesian coordinate system, the $\mathbf{e}_i$'s are the unit vectors $\hat{\mathbf{i}},\ \hat{\mathbf{j}},\ \hat{\mathbf{k}}$. The summation in (2.36) extends over only two terms because the electromagnetic wave is transverse, confining $\mathbf{E}$ to a plane normal to the direction of propagation [according to the coordinate convention we have selected, the $\mathbf{E}$ field is in the (x, y) plane].

The polarization can also be described in terms of basis vectors consisting of a right-circularly polarized component

$$\mathbf{E}_{\mathcal{R}} = E_{0\mathcal{R}}[\cos(\omega t - kz)\,\hat{\mathbf{i}} - \sin(\omega t - kz)\,\hat{\mathbf{j}}]$$

and a left-circularly polarized component

$$\mathbf{E}_{\mathcal{L}} = E_{0\mathcal{L}}[\cos(\omega t - kz)\,\hat{\mathbf{i}} + \sin(\omega t - kz)\,\hat{\mathbf{j}}].$$

An arbitrary elliptical polarization can then be written as

$$\begin{aligned} \mathbf{E} &= \mathbf{E}_{\mathcal{R}} + \mathbf{E}_{\mathcal{L}} \\ &= (E_{0\mathcal{R}} + E_{0\mathcal{L}}) \cos(\omega t - kz)\,\hat{\mathbf{i}} - (E_{0\mathcal{R}} - E_{0\mathcal{L}}) \sin(\omega t - kz)\,\hat{\mathbf{j}}. \end{aligned} \tag{2.38}$$

The geometrical construction that demonstrates the expression of an arbitrary elliptically polarized light wave in terms of right- and left-circularly polarized waves is shown in Figure 2.9. The use of circularly polarized waves as the basis set for describing polarization is discussed by Klein and Furtak [5].

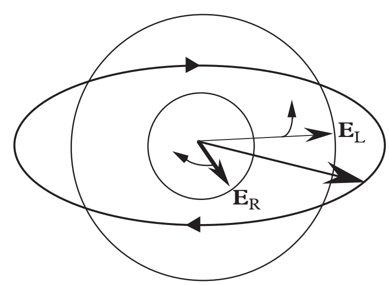

Figure 2.9 *Construction of elliptically polarized light from two circularly polarized waves.*

In the formalism associated with (2.37), the expansion coefficients a_i can be used to form a 2×2 matrix that in statistical mechanics is called the density matrix and in optics the coherency matrix [6]. The elements of this matrix are formed by the rule

$$\rho_{ij} = \mathbf{a}_i \mathbf{a}_j^*.$$

We will not develop the theory of polarization using the coherency matrix, but will simply use this matrix to justify the need for four independent measurements to characterize polarization. There is no unique set of measurements required by theory, but normally measurements made are of the *Stokes parameters*, which are directly related to the polarization ellipse of Figure 2.4. (We will see shortly that only three of the four measurements are independent. This will be in agreement with the definition of the coherency matrix, where $\rho_{ij} = \rho_{ij}^*$, i.e., the matrix is Hermitian.)

2.8.2 Stokes Parameters

The Stokes parameters (**Sir George Gabriel Stokes: 1819–1903**) of a light wave are measurable quantities, defined as follows:

- s_0 total flux density;
- s_1 difference between flux densities transmitted by a linear polarizer oriented parallel to the x-axis and one oriented parallel to the y-axis (the x- and y-axes are usually selected to be parallel to the horizontal and vertical directions in the laboratory);
- s_2 difference between flux densities transmitted by a linear polarizer oriented at 45° to the x-axis and one oriented at 135°;
- s_3 difference between flux densities transmitted by a right-circular polarizer and a left-circular polarizer.

The physical instruments that can be used to measure the parameters will be discussed in Chapter 13.

If the Stokes parameters are to characterize the polarization of a wave, they must be related to the parameters of the polarization ellipse. It is therefore important to establish that the Stokes parameters are variables of the polarization ellipse (2.33). In its current form, (2.33) contains no measurable quantities and thus must be modified if it is to be associated with the Stokes parameters. In the discussion of the Poynting vector, it was pointed out that the time average of the Poynting vector is the quantity observed when measurements are made of light waves. We must, therefore, find the time average of (2.33) if we wish to relate its parameters to observable quantities. To simplify the discussion, let us assume that the amplitudes of the orthogonally polarized waves E_{0x} and E_{0y} and their relative phase δ are constants. We will also use the shorthand notation for a time average introduced in (2.24):

$$\left\langle E_x^2 \right\rangle = \frac{1}{T} \int_{t_0}^{t_0+T} E_{0x}^2 [\cos(\omega t - kz)\cos\phi_1 - \sin(\omega t - kz)\sin\phi_1]^2 \, dt.$$

The time average of (2.33) can now be written as

$$\frac{\left\langle E_x^2 \right\rangle}{E_{0x}^2} + \frac{\left\langle E_y^2 \right\rangle}{E_{0y}^2} - 2\frac{\left\langle E_x E_y \right\rangle}{E_{0x}E_{0y}} \cos\delta = \sin^2\delta. \tag{2.39}$$

Multiplying both sides of (2.39) by $(2E_{0x}E_{0y})^2$ removes the terms in the denominators to give

$$4E_{0y}^2\left\langle E_x^2\right\rangle + 4E_{0x}^2\left\langle E_y^2\right\rangle - 8E_{0x}E_{0y}\left\langle E_xE_y\right\rangle\cos\delta = (2E_{0x}E_{0y}\sin\delta)^2.$$

The same argument that was used to simplify (2.25) can be used to obtain the time averages for the first two terms:

$$\left\langle E_x^2\right\rangle = \frac{E_{0x}^2}{2}, \qquad \left\langle E_y^2\right\rangle = \frac{E_{0y}^2}{2}.$$

The calculation of the time average in the third term,

$$\left\langle E_xE_y\right\rangle = \frac{1}{2}E_{0x}E_{0y}\cos\delta, \tag{2.40}$$

is left as Problem 2.12. With these time averages, (2.39) can be written as

$$4E_{0x}^2E_{0y}^2 - \left(2E_{0x}E_{0y}\cos\delta\right)^2 = \left(2E_{0x}E_{0y}\sin\delta\right)^2.$$

If $E_{0x}^2 + E_{0y}^2$ is added to both sides of this equation, it can be rewritten as

$$\left(E_{0x}^2 + E_{0y}^2\right)^2 - \left(E_{0x}^2 - E_{0y}^2\right)^2 - \left(2E_{0x}E_{0y}\cos\delta\right)^2 = \left(2E_{0x}E_{0y}\sin\delta\right)^2. \tag{2.41}$$

Each term in this equation can be identified with a Stokes parameter.

In our derivation we required that the amplitudes and relative phase of the two orthogonally polarized waves be constant, but we can relax this requirement and instead define the Stokes parameters as temporal averages. With this modification, the terms of (2.41) become

$$\begin{aligned} s_0 &= \left\langle E_{0x}^2\right\rangle + \left\langle E_{0y}^2\right\rangle, \quad s_1 = \left\langle E_{0x}^2\right\rangle - \left\langle E_{0y}^2\right\rangle, \\ s_2 &= \left\langle 2E_{0x}E_{0y}\cos\delta\right\rangle, \quad s_3 = \left\langle 2E_{0x}E_{0y}\sin\delta\right\rangle. \end{aligned} \tag{2.42}$$

Equation (2.41) can now be written as

$$s_0^2 - s_1^2 - s_2^2 = s_3^2. \tag{2.43}$$

For a polarized wave, only three of the Stokes parameters are independent. This agrees with the requirement placed upon elements of the Hermitian coherency matrix, introduced above.

With this demonstration of the connection between the Stokes parameters and the polarization ellipse, the Stokes parameters can be written in terms of the parameters of the polarization ellipse in Figure 2.4:

$$\begin{aligned} s_1 &= s_0\cos 2\varphi\cos 2\theta, \\ s_2 &= s_0\cos 2\varphi\sin 2\theta, \\ s_3 &= s_0\sin 2\varphi. \end{aligned} \tag{2.44}$$

It is this close relationship between the Stokes parameters and the polarization ellipse that makes the Stokes parameters a useful characterization of polarization.

The Stokes parameters can be used to define the *degree of polarization*

$$V = \frac{1}{s_0}\sqrt{s_1^2 + s_2^2 + s_3^2}. \tag{2.45}$$

[The equality (2.43) applies to completely polarized light, when $V = 1$.] The degree of polarization can be used to characterize any light source that is physically realizable. If the time averages in the definition of the Stokes parameters s_2 and s_3 in (2.42) are zero and if $\langle E_{0x}^2\rangle = \left\langle E_{0y}^2\right\rangle$, so that

$$s_0 = 2\left\langle E_{0x}^2\right\rangle, \quad s_1 = s_2 = s_3 = 0,$$

then the light wave is said to be unpolarized and $V = 0$.

Hans Mueller, a physics professor at MIT [6], pointed out that the Stokes parameters can be thought of as elements of a column matrix or 4-vector

$$\begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \end{bmatrix}.$$

This view will allow us to follow a polarized wave through a series of optical devices by using matrix algebra, as we will see later.

Poincaré's sphere (Henri Poincaré: 1854–1912) Before it was discovered that the Stokes parameters could be treated as elements of a column matrix, a geometrical construction was used to determine the effect of an anisotropic medium on polarized light. The parameters s_1, s_2, s_3 are viewed as the Cartesian coordinates of a point on a sphere of radius s_0. This sphere is called the *Poincaré sphere* [7] and is shown in Figure 2.10. On the sphere, right-hand polarized light is represented by points on the upper hemisphere. Linear polarization is represented by points on the equator. Circular polarization is represented by the poles.

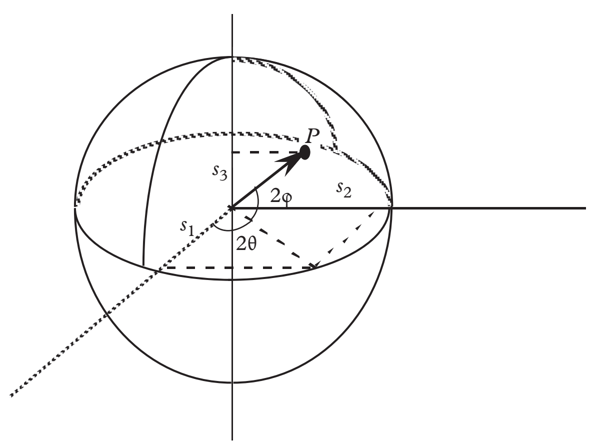

Figure 2.10 *Poincaré sphere.*

2.8.3 Jones Vector

There is one other representation of polarized light, complementary to the Stokes parameters, developed by **R. Clark Jones (1916–2004)** in 1941 and called the *Jones vector*. It is superior to the Stokes vector in that it handles light of a known phase and amplitude with a reduced number of parameters. It is inferior to the Stokes vector in that, unlike the Stokes representation, which is experimentally determined, the Jones representation cannot handle unpolarized or partially polarized light. The Jones vector is a theoretical construct that can only describe light with a well-defined phase and frequency. The density matrix formalism can be used to correct the shortcomings of the Jones vector, but then the simplicity of the Jones representation is lost.

It was shown earlier that if it is assumed that the coordinate system is such that the electromagnetic wave is propagating along the z-axis, then any polarization can be decomposed

Table 2.4 *Jones and Stokes vectors*

Jones vector	**Stokes vector**
Horizontal polarization	
$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$
Vertical polarization	
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$
+45° polarization	
$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$
–45° polarization	
$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$
Right-circular polarization	
$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$
Left-circular polarization	
$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$

into two orthogonal **E** vectors, say, for the purposes of this discussion, parallel to the x- and y-directions. The Jones vector is defined as a two-row column matrix consisting of the complex components in the x- and y-directions:

$$\mathbf{E} = \begin{bmatrix} E_{0x}e^{i(\omega t - \mathbf{k}\cdot\mathbf{r}+\phi_1)} \\ E_{0y}e^{i(\omega t - \mathbf{k}\cdot\mathbf{r}+\phi_2)} \end{bmatrix}. \tag{2.46}$$

If absolute phase is not an issue, then we may normalize this vector by dividing it by that number that simplifies the components but keeps the sum of the squared magnitudes of the components equal to one. For example, assuming that $E_{0x} = E_{0y}$, then

$$\mathbf{E} = E_{0x}e^{i(\omega t - \mathbf{k}\cdot\mathbf{r}+\phi_1)} \begin{bmatrix} 1 \\ e^{i\delta} \end{bmatrix}.$$

The normalized vector would comprise the terms contained within the square brackets, each divided by $\sqrt{2}$. The general form of the Jones vector is

$$\mathbf{E} = \begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}, \qquad \mathbf{E}^* = \begin{bmatrix} \mathbf{A}^* & \mathbf{B}^* \end{bmatrix}.$$

Some examples of Jones vectors and Stokes vectors are shown in Table 2.4.

2.9 Propagation in a Conducting Medium

In Chapter 1, we discussed the propagation of a wave with attenuation. In our discussion of the propagation of light, however, we have ensured that we would experience no loss by assuming $\sigma = 0$. We now relax that assumption and allow $\sigma \neq 0$. Maxwell's equations become

$$\nabla\cdot\mathbf{D} = 0, \qquad \nabla\cdot\mathbf{B} = 0,$$

$$\nabla\times\mathbf{H} = \mathbf{J} + \frac{\partial\,\mathbf{D}}{\partial t}, \qquad \nabla\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}.$$

We continue to neglect dynamic or resonant effects so that we may use the simple constitutive relations

$$\mathbf{J} = \sigma\mathbf{E}, \quad \mathbf{D} = \varepsilon\mathbf{E}, \quad \mathbf{B} = \mu\mathbf{H}.$$

where ε, μ, and σ are scalars, independent of time. Maxwell's equations in a medium with dissipation can be rewritten using these constitutive relations as

$$\begin{aligned} &\nabla\cdot\mathbf{E} = 0, && \nabla\cdot\mathbf{H} = 0, \\ &\nabla\times\mathbf{H} = \sigma\mathbf{E} + \varepsilon\frac{\partial\mathbf{E}}{\partial t}, && \nabla\times\mathbf{E} = -\mu\frac{\partial\mathbf{H}}{\partial t}. \end{aligned} \tag{2.47}$$

We now apply the same procedure used to derive the wave equation for free space,

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla \times \left(-\mu \frac{\partial \mathbf{H}}{\partial t}\right) = -\mu \frac{\partial}{\partial t}(\nabla \times \mathbf{H}),$$

$$-\mu \frac{\partial}{\partial t}(\nabla \times \mathbf{H}) = -\mu \frac{\partial}{\partial t}\left(\sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t}\right),$$

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}.$$

yielding the wave equation in a conducting medium:

$$\nabla^2 \mathbf{E} = \mu\sigma \frac{\partial \mathbf{E}}{\partial t} + \mu\varepsilon \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{2.48}$$

This wave equation is of the same form as (1.20). We can derive a similar equation for the magnetic field:

$$\nabla^2 \mathbf{B} = \mu\sigma \frac{\partial \mathbf{B}}{\partial t} + \mu\varepsilon \frac{\partial^2 \mathbf{B}}{\partial t^2}. \tag{2.49}$$

Equations (2.48) and (2.49) are called the *telegraph equations* and were developed by Oliver Heaviside (1850–1925). They are wave equations derived to explain the propagation of pulses on telegraph lines.

We see that the wave equation (2.48) contains a damping term $\partial \mathbf{E}/\partial t$ when we allow $\sigma \neq 0$. By comparing (2.48) and (1.20), we can see that the solution of (2.48) will be an electromagnetic wave that will experience attenuation proportional to $\mu\sigma$ as it propagates. Using (1.22) and (1.23), we can rewrite (2.47), for plane wave solutions, as

$$\begin{aligned} \nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t} \quad &\Rightarrow \quad \nabla \times \mathbf{E} = -i\omega\mu \mathbf{H}, \\ \nabla \times \mathbf{H} = \sigma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t} \quad &\Rightarrow \quad \nabla \times \mathbf{H} = i\omega\left(\varepsilon - \frac{i\sigma}{\omega}\right)\mathbf{E}. \end{aligned} \tag{2.50}$$

We can rewrite (2.48) in terms of these expressions for the curls of $\mathbf{E}$ and $\mathbf{H}$:

$$\nabla^2 \mathbf{E} + \omega^2 \mu \left(\varepsilon - \frac{i\sigma}{\omega}\right) \mathbf{E} = 0. \tag{2.51}$$

This has the form of the Helmholtz equation (1.13) if we replace k^2 in the latter by the complex function

$$\tilde{k}^2 = \omega^2 \mu \left(\varepsilon - i\frac{\sigma}{\omega}\right). \tag{2.52}$$

We use the identity

$$k = \frac{n\omega}{c} = \omega\sqrt{\mu\varepsilon},$$

to demonstrate that the equations for conducting media are identical to those derived for nonconducting media if the dielectric constant ε is replaced by a complex dielectric constant

$$\tilde{\varepsilon} = \varepsilon - i\left(\frac{\sigma}{\omega}\right). \tag{2.53}$$

This equation suggests that σ may contain a frequency dependence (in fact, in the c.g.s. system, the units of σ are s^{-1}; for copper, in c.g.s. units, $\sigma = 5.14 \times 10^{17}\ s^{-1}$). In condensed matter physics, one finds that the mobility of the electrons creates a frequency dependence that shows up in σ.

Since we have replaced k by the complex quantity

$$\tilde{k} = \omega\sqrt{\mu\left(\varepsilon - i\frac{\sigma}{\omega}\right)},$$

we must replace the index of refraction by a complex index. In the literature, this is accomplished in two ways:

$$\tilde{n} = n(1 - i\kappa), \tag{2.54}$$

$$\tilde{n} = n_1 - in_2.$$

We will use the notation in (2.54).

To find out how a plane wave propagates in this conductive medium, we simply replace the propagation constant k by

$$\tilde{k} = \tilde{n}\frac{\omega}{c} = \left(\frac{\omega n}{c}\right)(1 - i\kappa),$$

as we did with (1.24). κ is called the *extinction coefficient* and $n\kappa$ is called the *absorption coefficient*.

If we assume that $\mathbf{k}$ is parallel to the z-axis, then the plane wave is

$$\mathbf{E} = \mathbf{E}_0 e^{i\omega t} e^{-i\omega(n/c)(1-i\kappa)z},$$

$$\mathbf{E} = \mathbf{E}_0 e^{-(\omega n\kappa/c)z} e^{-i\omega(t-nz/c)}, \tag{2.55}$$

$$\mathcal{Re}\{\mathbf{E}\} = \mathbf{E}_0 e^{-(\omega n\kappa/c)z}\cos(\omega t - kz). \tag{2.56}$$

The wave described by (2.56) is a plane wave, attenuated by the exponential factor

$$e^{-(\omega/c)n\kappa z}. \tag{2.57}$$

Figure 2.11 *Blue laser light is shown propagating in xylene (above) and water (below). The water contains the dye rhodamine 6G in solution. The red rhodamine dye absorbs the blue light, and the beam rapidly decays to zero. Some of the energy absorbed by the dye is reemitted in the yellow to red region of the spectrum. This reemitted light caused the diffuse appearance of the light as it propagates in the water.*

Figure 2.11 displays the exponential decay of a light wave propagating in an absorbing medium. A layer of xylene floats upon water containing the dye rhodamine 6G in solution. The dye strongly absorbes a beam of blue light from a HeCd laser (440 nm). As can be seen in Figure 2.11, the blue light is rapidly attenuated once it enters the water. Some of the energy absorbed by the rhodamine is reemitted at longer wavelengths. The reemitted light travels in all directions, since it has no memory of the direction traveled by the blue light. For this reason, the beam of light in the water appears diffuse, and as can be seen, is orange in color.

To evaluate the absorption coefficient $n\kappa$ in terms of electromagnetic properties of the medium, we will derive a relationship between $n\kappa$ and σ. We rewrite (2.54) as

$$\tilde{n}^2 = n^2(1-\kappa^2-2i\kappa) = \frac{c^2}{\omega^2}\tilde{k}^2.$$

This can be used to express n^2 in terms of the constants of the material:

$$\tilde{n}^2 = c^2\mu\left(\varepsilon - i\frac{\sigma}{\omega}\right). \tag{2.58}$$

Equating real and imaginary terms, we obtain

$$n^2(1-\kappa^2) = c^2\mu\varepsilon, \quad 2n^2\kappa = c^2\frac{\mu\sigma}{\omega}.$$

We can use these two relationships to find

$$n^2 = \frac{c^2}{2}\left[\sqrt{\mu^2\varepsilon^2 + \left(\frac{\mu\sigma}{\omega}\right)^2} + \mu\varepsilon\right], \tag{2.59}$$

$$n^2\kappa^2 = \frac{c^2}{2}\left[\sqrt{\mu^2\varepsilon^2 + \left(\frac{\mu\sigma}{\omega}\right)^2} - \mu\varepsilon\right]. \tag{2.60}$$

Note that when $\sigma = 0$, $\kappa = 0$, and we obtain the free-space result, (2.11):

$$n^2 = \frac{\mu\varepsilon}{\mu_0\varepsilon_0}.$$

An estimate of the magnitude of the quantities under the radicals in (2.59) and (2.60) can be obtained by using values for copper, where, in MKS units, $\sigma = 5.8 \times 10^7$ mho/m and $n = 0.62$ at $\lambda = 589.3$ nm. (The index of refraction is less than one, which implies that the phase velocity is greater than the speed of light. This apparent contradiction of a fundamental postulate of the theory of relativity will be discussed during the study of dispersion in Chapter 7.) The two terms under the radical are

$$\frac{\mu\sigma}{\omega} = \frac{(4\pi \times 10^{-7})(5.8 \times 10^7)(5.893 \times 10^{-7})}{(2\pi)(3 \times 10^8)} = 2.3 \times 10^{-14}\ \text{s}^2/\text{m}^2,$$

$$\mu\varepsilon = \mu_0\varepsilon_0 n^2 = (4\pi \times 10^{-7})(8.8542 \times 10^{-12})(0.62)^2 = 4.3 \times 10^{-18}\ \text{s}^2/\text{m}^2.$$

By comparing the relative magnitude of these two terms, we are justified in assuming that $\sigma/\omega \gg \varepsilon$ and can make the approximation

$$n^2\kappa^2 \approx \frac{c^2\mu\sigma}{2\omega},$$

$$n\kappa = c\sqrt{\frac{\mu\sigma}{2\omega}}. \tag{2.61}$$

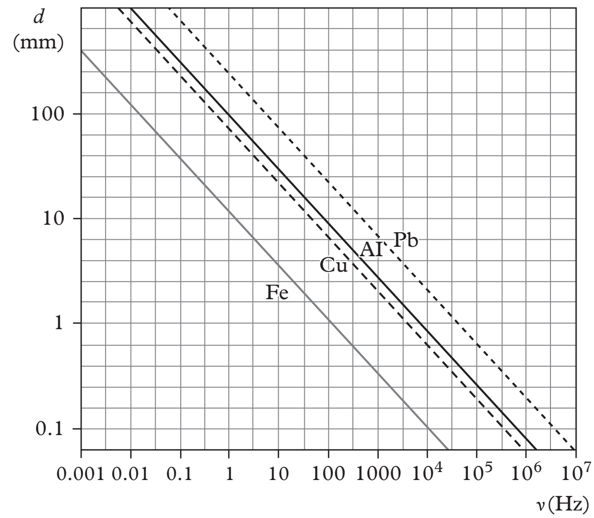

Figure 2.12 *Skin depth for a few metals.*

We use (2.61) to find the depth at which an electromagnetic wave is attenuated, to $1/e$ of its original energy, when propagating into a conductor. At that depth, denoted by d, the exponent in (2.57) will equal 1, and thus

$$\frac{\omega}{c} n\kappa d = \frac{2\pi}{\lambda_0} n\kappa d = 1,$$

$$d = \frac{\lambda_0}{2\pi n\kappa} \approx \frac{\lambda_0}{2\pi c}\sqrt{\frac{2\omega}{\mu\sigma}},$$

$$d = \sqrt{\frac{2}{\mu\sigma\omega}}. \tag{2.62}$$

The depth d is called the *skin depth*. Figure 2.12 shows the dependence of skin depth on frequency for some metals.

2.10 Summary

In this chapter, Maxwell's equations were used to obtain the wave equation for free space,

$$\nabla^2\mathbf{E} = \mu\varepsilon\frac{\partial^2\,\mathbf{E}}{\partial t^2}.$$

and for a conductive medium,

$$\nabla^2\mathbf{E} = \mu\sigma\frac{\partial\mathbf{E}}{\partial t} + \mu\varepsilon\frac{\partial^2\mathbf{E}}{\partial t^2}.$$

While the forms of these two equations appear quite different, we demonstrated that plane wave solutions existed for both equations when the dielectric constant of the conductive medium was replaced by a complex constant

$$\tilde{\varepsilon} = \varepsilon - i\left(\frac{\sigma}{\omega}\right).$$

This replacement means that the optical properties of a conductive material are described by a complex index of refraction

$$\tilde{n} = n(1 - i\kappa).$$

A plane wave propagating in conductive material has an amplitude attenuated by the exponential factor

$$e^{-(\omega/c)n\kappa z}.$$

By manipulation of Maxwell's equations, we were able to show that the propagation velocity of a light wave is governed by the electrical properties of the medium. The index of refraction was used to indicate the propagation velocity in the medium, relative to the propagation velocity in a vacuum:

$$n = \frac{c}{v} = \sqrt{\frac{\varepsilon\mu}{\varepsilon_0\mu_0}}.$$

Also, we were able to demonstrate that light waves must be transverse waves and that the magnitudes of the electric and magnetic fields are related by

$$n|\mathbf{E}| = c|\mathbf{B}|.$$

By comparing the forces experienced by a charged particle in an electromagnetic field, we found that we could describe the polarization of an electromagnetic wave by the electric field vector. We developed the formalism necessary for discussion of polarization, but delayed a discussion of the manipulation of a light wave's polarization until Chapter 13.

2.11 Problems

2.1. Light is traveling in glass ($n = 1.5$). If the amplitude of the electric field of the light is 100 V/m, what is the amplitude of the magnetic field? What is the magnitude of the Poynting vector?

2.2. A 60 W monochromatic point source is radiating equally in all directions in a vacuum. What is the electric field amplitude 2 m from the source?

2.3. The flux density at the Earth's surface due to sunlight is $I = 1.34 \times 10^3$ J/(m^2·s). Calculate the electric and magnetic fields at the Earth's surface by assuming that the average Poynting vector is equal to this flux density.

2.4. What is the flux density of light needed to keep a glass sphere of mass 10^{-8} g and diameter 2×10^{-5} m floating in midair?

2.5. An 85 kg astronaut has only a flashlight to propel him in space. If the flashlight emits 1 W of light in a parallel beam for one hour, how fast will the astronaut be going at the end of the hour, assuming he started at rest?

2.6. What are the polarizations of the following waves:

$$\mathbf{E} = E_0\left[\cos(\omega t - kz)\,\hat{\mathbf{i}} + \cos\left(\omega t - kz + \frac{5}{4}\pi\right)\hat{\mathbf{j}}\right],$$

$$\mathbf{E} = E_0\left[\cos(\omega t + kz)\,\hat{\mathbf{i}} + \cos\left(\omega t + kz - \frac{1}{4}\pi\right)\hat{\mathbf{j}}\right],$$

$$\mathbf{E} = E_0\left[\cos(\omega t - kz)\,\hat{\mathbf{i}} - \cos\left(\omega t - kz + \frac{1}{6}\pi\right)\hat{\mathbf{j}}\right]?$$

2.7. Show that the addition of two elliptically polarized waves propagating along the z-axis results in another elliptically polarized wave.

2.8. Write an expression, in MKS units, for a plane electromagnetic wave, with a wavelength of 500 nm and an intensity of 53.2 W/m^2, propagating in the z-direction. Assume that the wave is linearly polarized at an angle of 45° to the x-axis.

2.9. Using conventional vector notation, prove that a right- and a left-circularly polarized wave can be combined to yield a linearly polarized wave. Carry out the same demonstration using the Jones vector notation. What requirement must be placed on the two circularly polarized waves? Sketch the geometrical construction that demonstrates the combination of circularly polarized waves to generate a linearly polarized wave.

2.10. Write the equation for a plane wave propagating in the positive z-direction that has right elliptical polarization with the major axis of the ellipse parallel to the x-axis. Use both the conventional vector and the Jones vector notation.

2.11. Describe the polarization of a wave with the Jones vector

$$\begin{bmatrix} -i \\ 2 \end{bmatrix}.$$

Write the Jones vector that is orthogonal to this vector and describe its polarization.

2.12. Prove that (2.42) is correct using the expressions

$$E_x = E_{0x}[\cos(\omega t - kz)\cos\phi_1 - \sin(\omega t - kz)\sin\phi_1],$$

$$E_y = E_{0y}[\cos(\omega t - kz)\cos\phi_2 - \sin(\omega t - kz)\sin\phi_2]$$

for the two orthogonally polarized electric fields.

2.13. Demonstrate, using the Jones vector notation, that right- and left-circularly polarized light waves are orthogonal.

2.14. Find the skin depth for seawater with a resistivity $\rho = 0.20\ \Omega$/m for $\nu = 30$ kHz and 30 MHz. What frequency should we use to communicate with a submarine that will not be deeper than 100 m?

2.15. At what frequency would the approximation used to obtain (2.26) produce a 10% error?

2.16. If a 1 kW laser beam is focused to a spot with an area of 10^{-9} m^2, what is the amplitude of the electric field at the focus?

2.17. The human eye is sensitive to light of wavelengths from approximately 600 nm (red) to 400 nm (blue). (a) Calculate the frequency of both wavelengths. (b) Find the energy of the photons associated with the red and blue wavelength limits.

2.18. If green light, 500 nm, could be frequency modulated to 0.1% of the light wave's frequency, calculate the number of 6 MHz bandwidth TV channels that could be carried by the modulation.

2.19. Given the Stokes vector

$$\begin{bmatrix} 1 \\ 0 \\ -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix},$$

(a) calculate the degree of polarization, (b) determine the orthogonal vector, and (c) draw the polarization ellipse.

2.20. How thin must a sheet of iron be if it is one skin depth thick? How many atoms thick is such a sheet?

REFERENCES

1. Maxwell, J.C., *A dynamical theory of the electromagnetic field*. Phil. Trans. R. Soc. Lond., 1865. **155**: p. 459–512.
2. Feynman, R.P., *QED, The Strange Theory of Light and Matter*. 2006: Princeton University Press: p. 192.
3. Wangsness, R.K., *Electromagnetic Fields*. 2nd ed. 1986: Wiley.
4. McMaster, W.H., *Polarization and the Stokes parameters*. Am. J. Phys., 1954. **22**: p. 351–362.
5. Klein, M.V. and T.E. Furtak, *Optics*. 2nd ed. 1986: Wiley.
6. Mueller, H., *The foundation of optics*. J. Opt. Soc. Am, 1948. **38**: p. 661.
7. Poincaré, H., *Théorie mathématique de la lumiére*. Vol. 2. 1892: Georges Carré.

Appendix 2A Vectors

We will review a few properties of vectors that will be of use in our discussion of light. A vector is a quantity with both magnitude and direction; it can be defined in terms of unit vectors along the three orthogonal axes of a Cartesian coordinate system:

$$\mathbf{E} = E_x\hat{\mathbf{i}} + E_y\hat{\mathbf{j}} + E_z\hat{\mathbf{k}}.$$

To add two vectors, we add like components:

$$\mathbf{E}_1 + \mathbf{E}_2 = (E_{1x} + E_{2x})\hat{\mathbf{i}} + (E_{1y} + E_{2y})\hat{\mathbf{j}} + (E_{1z} + E_{2z})\hat{\mathbf{k}}.$$

2A.1 Products

There are two ways to multiply vectors:

- The *scalar (dot) product* is given by

$$\mathbf{E} \cdot \mathbf{H} = EH\cos\theta = E_xH_x + E_yH_y + E_zH_z, \tag{2A.1}$$

  where θ is the angle between **E** and **H**. This product is a scalar quantity that gives the projection of one vector onto the second vector. If **E** and **H** are perpendicular, then their dot product is zero.

- The *vector (cross) product* is given by

$$\mathbf{E} \times \mathbf{H} = EH\sin\theta\,\hat{\mathbf{n}},$$

  where $\hat{\mathbf{n}}$ is a unit vector normal to the plane formed by **E** and **H**. The cross product is a vector with a magnitude equal to the area of the parallelogram formed by **E** and **H**; it is zero if the two vectors are parallel. The components of this vector are calculated as follows:

$$\begin{aligned}\mathbf{E} \times \mathbf{H} &= \begin{bmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ E_x & E_y & E_z \\ H_x & H_y & H_z \end{bmatrix} \\ &= (E_yH_z - E_zH_y)\hat{\mathbf{i}} - (E_xH_z - E_zH_x)\hat{\mathbf{j}} + (E_xH_y - E_yH_x)\hat{\mathbf{k}}.\end{aligned} \tag{2A.2}$$

  We will find use for the vector triple product relationship

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}. \tag{2A.3}$$

2A.2 Derivatives

A vector operator called the *del operator* is defined in a Cartesian coordinate system as

$$\nabla = \hat{\mathbf{i}}\frac{\partial}{\partial x} + \hat{\mathbf{j}}\frac{\partial}{\partial y} + \hat{\mathbf{k}}\frac{\partial}{\partial z}. \tag{2A.4}$$

We can treat this operator as a vector and calculate three products of use in optics:

- The *gradient* of a scalar:

$$\nabla V = \frac{\partial V}{\partial x}\hat{\mathbf{i}} + \frac{\partial V}{\partial y}\hat{\mathbf{j}} + \frac{\partial V}{\partial z}\hat{\mathbf{k}}. \tag{2A.5}$$

  The gradient is a vector giving the magnitude and direction of the fastest rate of change of the scalar quantity V.

- The *divergence* of a vector:

$$\nabla \cdot \mathbf{E} = \frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z}. \tag{2A.6}$$

The divergence gives the amount of flux flowing toward (negative) or away from (positive) a point. If the divergence is zero, then there are no sources or sinks in the volume.

- The *curl* of a vector:

$$\nabla \times \mathbf{E} = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right)\hat{\mathbf{i}} + \left(\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}\right)\hat{\mathbf{j}} + \left(\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y}\right)\hat{\mathbf{k}}. \tag{2A.7}$$

A physical interpretation of this operation can be made easily if the vector is a velocity, then when the curl of the velocity is nonzero, rotation is also occurring.

If we calculate the divergence of a gradient of a scalar, we obtain a scalar function, the *Laplacian*:

$$\nabla \cdot \nabla V = \nabla^2 V = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2}. \tag{2A.8}$$

For a vector quantity, we have

$$\nabla^2 \mathbf{E} = \nabla^2 E_x \,\hat{\mathbf{i}} + \nabla^2 E_y \,\hat{\mathbf{j}} + \nabla^2 E_z \,\hat{\mathbf{k}}. \tag{2A.9}$$

Table 2A.1 *Operations with the del operator in spherical coordinates (r, θ, ϕ), where subscripts r, θ, and ϕ denote the respective components*

$(\nabla V)_r = \dfrac{\partial V}{\partial r}$	$(\nabla V)_\theta = \dfrac{1}{r}\left(\dfrac{\partial V}{\partial \theta}\right)$	$(\nabla V)_\phi = \dfrac{1}{r\sin\theta}\left(\dfrac{\partial V}{\partial \phi}\right)$
$\nabla \cdot \mathbf{E} = \dfrac{1}{r}\left[\dfrac{\partial (rE_r)}{\partial r}\right] + \dfrac{1}{r\sin\theta}\left[\dfrac{\partial (E_\theta \sin\theta)}{\partial \theta}\right] + \dfrac{1}{r\sin\theta}\left(\dfrac{\partial E_\phi}{\partial \phi}\right)$		
$\nabla^2 V = \dfrac{1}{r^2}\dfrac{\partial}{\partial r}\left(r^2 \dfrac{\partial V}{\partial r}\right) + \dfrac{1}{r^2 \sin\theta}\dfrac{\partial}{\partial \theta}\left(\sin\theta \dfrac{\partial V}{\partial \theta}\right) + \dfrac{1}{r^2 \sin^2\theta}\dfrac{\partial^2 V}{\partial \phi^2}$		
$(\nabla \times \mathbf{E})_r = \dfrac{1}{r\sin\theta}\left[\dfrac{\partial}{\partial \theta}(E_\phi \sin\theta) - \dfrac{\partial E_\theta}{\partial \phi}\right]$	$(\nabla \times \mathbf{E})_\theta = \dfrac{1}{r}\left[\dfrac{1}{\sin\theta}\dfrac{\partial E_r}{\partial \phi} - \dfrac{\partial}{\partial r}(rE_\phi)\right]$	$(\nabla \times \mathbf{E})_\phi = \dfrac{1}{r}\left[\dfrac{\partial}{\partial r}(rE_\theta) - \dfrac{\partial E_r}{\partial \theta}\right]$
$(\nabla^2 \mathbf{E})_r = \nabla^2 E_r - \dfrac{2E_r}{r^2} - \dfrac{2}{r^2 \sin\theta}\dfrac{\partial (E_\theta \sin\theta)}{\partial \theta} - \dfrac{2}{r^2 \sin\theta}\dfrac{\partial E_\phi}{\partial \phi}$	$(\nabla^2 \mathbf{E})_\theta = \nabla^2 E_\theta + \dfrac{2}{r^2}\dfrac{\partial E_r}{\partial \theta} - \dfrac{E_\theta}{r^2 \sin^2\theta} - \dfrac{2\cos\theta}{r^2 \sin^2\theta}\dfrac{\partial E_\phi}{\partial \phi}$	$(\nabla^2 \mathbf{E})_\phi = \nabla^2 E_\phi + \dfrac{2}{r^2 \sin\theta}\dfrac{\partial E_r}{\partial \phi} + \dfrac{2\cos\theta}{r^2 \sin^2\theta}\dfrac{\partial E_\theta}{\partial \phi} - \dfrac{E_\phi}{r^2 \sin^2\theta}$

Several other identities involving the del operator will come in handy in our study of optics:

$$\nabla \times \nabla V = 0, \tag{2A.10}$$

$$\nabla \cdot \nabla \times \mathbf{E} = 0, \tag{2A.11}$$

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}, \tag{2A.12}$$

$$\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{A} \cdot (\nabla \times \mathbf{B}) + \mathbf{B} \cdot (\nabla \times \mathbf{A}). \tag{2A.13}$$

We have displayed all of the above relations in a Cartesian coordinate system; similar expressions can be derived using spherical coordinates. We list them in Table 2A.1 for the reader's convenience.

Appendix 2B Electromagnetic Units

Of all topics in physics, perhaps the one that introduces the most confusion is the subject of electromagnetic units. In this book, we use MKS units, but the reader will find most older optics books will use c.g.s. units.

Table 2B.1 *Electromagnetic units*

c.g.s.	Units	MKS	Units
c	cm/s	$\dfrac{1}{\sqrt{\mu_0 \varepsilon_0}}$	m/s
$\mathbf{D}$	$12\pi \times 10^{-5}$ statcoulomb/cm^2	$\sqrt{4\pi\varepsilon_0}\ \mathbf{D}$	coulomb/m
$\mathbf{B}$	10^4 gauss	$\sqrt{\dfrac{4\pi}{\mu_0}}\ \mathbf{B}$	weber/m^2
$\mathbf{H}$	$4\pi \times 10^{-3}$ oersted	$\sqrt{4\pi\mu_0}\ \mathbf{H}$	amp-turn/m
$\mathbf{E}$	$\dfrac{1}{3} \times 10^{-4}$ statcoulomb/cm	$\sqrt{4\pi\varepsilon_0}\ \mathbf{E}$	volt/m
$\mathbf{J}$	3×10^5 statamp/cm^2	$\dfrac{\mathbf{J}}{\sqrt{4\pi\varepsilon_0}}$	amp/m^2
σ	9×10^9 s^{-1}	$\dfrac{\sigma}{4\pi\varepsilon_0}$	mho/m
ρ	3×10^3 statcoulomb/cm^3	$\dfrac{\rho}{\sqrt{4\pi\varepsilon_0}}$	coulomb/m^3
ε		$\dfrac{\varepsilon}{\varepsilon_0}$	
μ		$\dfrac{\mu}{\mu_0}$	

In rationalized MKS units, ε and μ in a vacuum have the following values

$$\begin{aligned}\varepsilon_0 &= 8.8542 \times 10^{-12}\ \mathrm{C/(N \cdot m^2)}\ \ (= \mathrm{F/m})\\ &\approx \left(\frac{1}{36\pi}\right) \times 10^{-9}\ \mathrm{F/m},\\ \mu_0 &= 4\pi \times 10^{-7} \mathrm{N \cdot s^2/C^2}\ \ (= \mathrm{H/m}).\end{aligned}$$

In c.g.s. units, $\varepsilon_0 = \mu_0 = 1$.

Maxwell's equations in c.g.s. units are written as

$$\begin{aligned}\nabla \cdot \mathbf{D} &= 4\pi\rho,\\ \nabla \cdot \mathbf{B} &= 0,\\ \nabla \times \mathbf{H} &= \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t},\\ \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} &= 0.\end{aligned}$$

A good discussion of the subject of units can be found in Jackson [1]. Here, we provide a brief table to aid the reader in converting from one system to the other.

REFERENCE

1. Jackson, J.D., *Clasical Electrodynamics*. 3rd ed. 1998: Wiley.