1

Introduction

Then we shall rise and view ourselves with clearer eyes.

Henry King, bishop of Chichester (1592–1669)

1.1 Scientific rationale

All models are wrong, but some are useful.

George E. P. Box

The subject of this book is remote sensing, that is, seeing "with clearer eyes." In particular, it is concerned with how light is emitted and scattered by media composed of discrete particles and what can be learned about such a medium from its scattering properties.

If you stop reading now and look around, you will notice that most of the surfaces you see consist of particulate materials. Sometimes the particles are loose, as in soils or clouds. Sometimes they are embedded in a transparent matrix, as in paint, which consists of white particles in a colored binder. Or they may be fused together, as in rocks, or tiles which consist of sintered ceramic powder. Even vegetation is a kind of particulate medium in which the "particles" are leaves and stems. These examples show that if we wish to quantitatively interpret the electromagnetic radiation that reaches us, rather than simply form an image from it, it is necessary to consider the scattering and propagation of light within nonuniform media.

One of the first persons to use remote sensing to learn about the surface of a planet was Galileo Galilei. Galileo (1638) noticed that the full Moon, as it rose over his garden wall opposite the setting Sun, was darker than the sunlit wall. He also noted that the bidirectional reflectance function (of course, he did not use that term) of the lunar surface was diffuse in nature, rather than specular. From those observations he argued that the Moon was not a smooth, perfectly reflecting, crystalline sphere, as the prevailing wisdom of the day supposed, but was a planet not unlike the Earth.

In the intervening years since Galileo, remote sensing of planetary surfaces has become a quantitative science that has been used by two major groups of scientists. One is the community of planetary scientists. Virtually everything we know about the surfaces of the other bodies of the solar system comes to us through remote sensing. Most spacecraft missions are flybys or orbiters, and even landers, manned or unmanned, can sample only tiny portions of the surfaces of the bodies on which they set down. The second group consists of those scientists who are concerned with processes on the surface of the Earth, including geologists, meteorologists, geographers, and agronomists, who use remote sensing to study the Earth from balloons, aircraft, and satellites. However, the book is sufficiently general that it should be useful to any scientist or engineer who uses reflectance or emittance as an analytical tool.

In addition, reflectance spectroscopy turns out to be a powerful method for quantitatively measuring the characteristic absorption spectrum of a material. It has several advantages over more conventional methods. The dynamic range of the technique is large, typically four or more orders of magnitude in the absorption coefficient. The method is effective when the imaginary component of the refractive index is in the range of 10^{-3} to 10^{-1}, where both transmission and specular reflection techniques are difficult. Finally, sample preparation is convenient and simply requires grinding and sieving the material. Thus, this book should also be of interest to anyone who measures absorption coefficients.

The purpose of this book is to present quantitative models that describe the diffuse scattering and thermal emission of electromagnetic radiation from particulate media, such as planetary regoliths or powders in the laboratory. There are two general approaches to this problem. The first is to start from Maxwell's electromagnetic equations and attempt to find exact solutions. However, even with the use of modern high-speed compters this turns out to be impossible except for systems that are so simple as to have little resemblance to actual media, and the amount of computer time required is so long as to be impractical for most applications. Such solutions are useful primarily for illuminating some of the physical processes involved in light scattering by particulate media. However, at the present stage of our computational abilities they cannot justify the rather astonishing claim made by some persons that these processes are understood perfectly.

The second method, and the one used in this book, is based primarily on the equation of radiative transfer. This infamous equation is so notoriously intractable that most students have tried to stay as far away from it as possible. However, it will be seen that, by making appropriate approximations, solutions that are mathematically simple but surprisingly accurate can be obtained. (For example, see Figure 8.10.) Essentially, the radiative transfer equation assumes that photons of light can be treated as a gas diffusing between the particles of the medium. Even though the mathematical basis of this assumption has not been fully established, and some of the solutions presented in this book are not exact, this approach can be justified on several grounds:

(1) In most remote-sensing measurements the accuracy with which *absolute* reflectances can be measured is usually relatively low, typical uncertainties being of the order of $\pm 10\%$. Even in the laboratory, considerable effort must be expended to improve this figure substantially.
(2) Most applications deal with large, irregular particles whose scattering properties are poorly known, so that a detailed numerical calculation may give a misleading impression of high accuracy.
(3) Often in applications to planetary science a first-order estimate is sufficient to explain an observation or evaluate a hypothesis, and greater precision is unwarranted by the data.
(4) To evaluate the effect of varying a parameter, a numerical solution must be repeated many times, whereas if an analytic expression is available the effect can often be ascertained by inspection.

Thus, in most cases, exact numerical solutions are no more useful and are much less convenient than approximate analytic ones. The philosophy of this book is to be as rigorous as possible but, where necessary, to make approximations that retain the essential physics and at the same time are sufficiently simple that solutions to the problems of interest can be given by closed analytic functions.

1.2 About this book

The book is aimed at advanced undergraduate and beginning graduate students in the physical sciences. It is assumed that the reader has had at least an introductory course in physics, plus calculus through differential equations, and has a slight familiarity with vectors and complex variables. For those whose knowledge of mathematics is limited or rusty, brief reviews of those aspects of vector calculus, complex variables, and standard solutions of the wave equation that will be needed in this book have been included in Appendixes A–C.

I have tried to use notations that are in wide use, but this is not always possible in a treatise that covers a large range of topics. A list of symbols is included as Appendix E. Occasionally I have had to use the same symbol for two different quantities; in that case the meaning should be obvious from the context.

The book is divided into four parts. The first part, consisting of Chapters 2–4, introduces the general subjects of wave propagation and absorption in continuous media, polarization, and specular reflection from boundaries. The second part, Chapters 5 and 6, describes the scattering of light by single particles. The third part, Chapters 7–14, treats diffuse reflectance and polarization by particulate media and applications to reflectance spectroscopy. Thermal emittance is treated in Chapters 15 and 16. The discussion of thermal emittance spectroscopy can be brief because Kirchhoff's laws show that the processes of emittance and reflectance are complementary. Hence, the results of the earlier chapters on reflectance can be carried over directly to emittance.

Finally, an important warning must be issued to those who might want to use equations printed in this book to interpret reflectance or emittance data. I have tried to ensure that the important equations are correct, but I doubt that I have been completely successful. All practitioners of science know that Murphy's law is real, and in the context of this book the law can be phrased, "The equation you need the most contains a typographical error." I have tried to give sufficient detail that a reader can readily reproduce most of the derivations. Always rederive a critical equation!

2

Electromagnetic wave propagation

2.1 Maxwell's equations

Because reflectance spectroscopy uses electromagnetic radiation to probe matter, this book begins with Maxwell's electromagnetic equations and the solutions to them that describe propagating plane waves. For those whose knowledge of vectors may be a bit rusty, a brief review of vector notation is provided in Appendix A.1.

In their general form, Maxwell's equations can be written as follows:

$$\text{div } \mathbf{D}_e = \rho_e, \tag{2.1}$$

$$\text{div } \mathbf{B}_m = 0, \tag{2.2}$$

$$\text{curl } \mathbf{E}_e = -\partial \mathbf{B}_m / \partial t, \tag{2.3}$$

$$\text{curl } \mathbf{H}_m = \mathbf{j}_e + \partial \mathbf{D}_e / \partial t. \tag{2.4}$$

In these equations, $\mathbf{E}_e$ is the electric field, $\mathbf{D}_e$ is the electric displacement, $\mathbf{B}_m$ is the magnetic-induction field, $\mathbf{H}_m$ is the magnetic intensity, ρ_e is the electric-charge density, $\mathbf{j}_e$ is the electric current density, t is the time, and "div" and "curl" are, respectively, the vector divergence and curl operators. The reader is referred to the many excellent textbooks on electromagnetic theory for detailed derivations and more rigorous discussions of these equations, including Stratton (1941), Panofsky and Phillips (1962), Marion (1965), Elliott (1966), Landau and Lifschitz (1975), and Jackson (1999).

Equation (2.1) states that electric charges can generate electric fields and that the field lines diverge from or converge toward the charges. Equation (2.2) states that there are no sources of magnetic fields that are analogous to electric charges; that is, magnetic monopoles do not exist. According to equation (2.3), electric fields can also be generated by magnetic fields that change with time, and electric fields generated in this manner tend to coil or curl around the magnetic-field lines. Similarly, according to equation (2.4), magnetic fields can be generated by both electric currents and time-varying electric fields, and the magnetic lines of force tend to curl around these sources.

Equations (2.1) – (2.4) are called the *field equations*. In order to solve them, additional relations, known as *constitutive equations*, that connect two or more of the quantities in the field equations are needed, along with appropriate boundary conditions. The constitutive relations describe the way that microscopic charges and currents generated in the medium by the applied field alter those fields. In the remainder of this chapter, solutions of the field equations that describe propagating plane waves will be discussed for several different types of constitutive equations.

2.2 Electromagnetic waves in free space

2.2.1 The wave equation

The simplest case of electromagnetic radiation is wave propagation in a vacuum. In free space there are no charges or currents, so the constitutive equations have the simple form

$$\rho_e = 0, \tag{2.5}$$

$$\mathbf{j}_e = 0, \tag{2.6}$$

$$\mathbf{D}_e = \varepsilon_{e0}\mathbf{E}_e, \tag{2.7}$$

$$\mathbf{B}_m = \mu_{m0}\mathbf{H}_m, \tag{2.8}$$

where ε_{e0} is the permittivity of free space ($\varepsilon_{e0} = 8.85 \times 10^{-12}$ coulomb^2/newton-meter^2), and μ_{m0} is the permeability of free space (($\mu_{m0} = 12.57 \times 10^{-7}$ amperes/m^2). Then the field equations become

$$\text{div } \mathbf{E}_e = 0, \tag{2.9}$$

$$\text{div } \mathbf{B}_m = 0, \tag{2.10}$$

$$\text{curl } \mathbf{E}_e = -\partial \mathbf{B}_m/\partial t, \tag{2.11}$$

$$\text{curl } \mathbf{B}_m = \mu_{m0}\varepsilon_{e0}\partial \mathbf{E}_e/\partial t. \tag{2.12}$$

These equations may be combined to yield relations that each contain only one field quantity, as follows. Take the curl of both sides of (2.11) and use vector identity (A.9) from Appendix A:

$$\text{curl}(\text{curl}\,\mathbf{E}_e) = \text{grad}(\text{div}\,\mathbf{E}_e) - (\text{div}\cdot\text{grad})\mathbf{E}_e = -\partial(\text{curl}\,\mathbf{B}_m)/\partial t. \tag{2.13}$$

Using (2.9) and (2.12) in (2.13) gives

$$\nabla^2\mathbf{E}_e - \mu_{m0}\varepsilon_{e0}\partial^2\mathbf{E}_e/\partial t^2 = 0, \tag{2.14}$$

where $\nabla^2 = \text{div} \cdot \text{grad}$ is the Laplacian operator (often called "del-squared"). Applying a similar analysis starting with the magnetic field, equation (2.12), yields

$$\nabla^2 \mathbf{B}_m - \mu_{m0}\varepsilon_{e0}\partial^2 \mathbf{B}_m/\partial t^2 = 0. \tag{2.15}$$

Equations (2.13) and (2.14) are of the general form

$$\nabla^2 f = (1/\upsilon^2)\partial^2 f/\partial t^2. \tag{2.16}$$

Equation (2.16) is known as the wave equation and describes a disturbance at position $\boldsymbol{r}$ of arbitrary shape propagating with velocity υ. If the geometry of the situation has plane-parallel symmetry, the general solution of (2.16) is $f = f(K\varnothing)$, where K is any constant, $\varnothing = \mathbf{u}_p \cdot \mathbf{r} \pm \upsilon t$ is the phase, $\mathbf{u}_p$ is a unit vector parallel to the direction of propagation, and f is an arbitrary function. All points satisfying $\mathbf{u}_p \cdot \mathbf{r} =$ constant lie on a plane perpendicular to $\mathbf{u}_p$ and passing through the point $\mathbf{r}$ (Figure 2.1).

In the remainder of this chapter, only plane waves will be considered. Without loss of generality, the direction of propagation may be taken to lie along the z-axis. In this case, f is independent of x and y, the differential vector operator "del" is $\nabla = \mathbf{u}_z \partial/\partial z$, where $\mathbf{u}_z$ is a unit vector pointing in the z direction, and $\varnothing = z \pm \upsilon t$. Then (2.16) becomes

$$\partial^2 f/\partial z^2 - (1/\upsilon^2)\partial^2 f/\partial t^2 = 0. \tag{2.17}$$

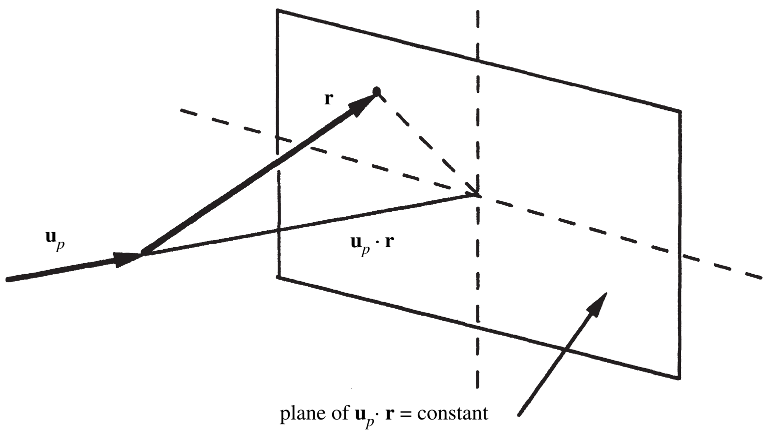

Figure 2.1

The proof that this equation is satisfied by any arbitrary function $f(K\varnothing)$ is almost trivial:

$$\frac{\partial^2 f}{\partial z^2} = \frac{d}{d\varnothing}\left(\frac{df}{d\varnothing}\frac{\partial\varnothing}{\partial z}\right)\frac{\partial\varnothing}{\partial z} = K^2\frac{d^2 f}{d\varnothing^2}, \tag{2.18}$$

$$\frac{1}{\upsilon^2}\frac{\partial^2 f}{\partial t^2} = \frac{1}{\upsilon^2}\frac{d}{d\varnothing}\left(\frac{df}{d\varnothing}\frac{\partial\varnothing}{\partial t}\right)\frac{\partial\varnothing}{\partial t} = K^2\frac{d^2 f}{d\varnothing^2}. \tag{2.19}$$

If $\varnothing = z - \upsilon t$, f represents a pattern whose shape is described by $f(K\varnothing)$ moving in the positive z direction with velocity υ; if $\varnothing = z + \upsilon t$, the pattern is moving toward negative z.

If the propagating pattern is periodic and repeats over a distance λ, then at any point z as the wave passes its pattern will repeat over a time interval $t = \lambda/\upsilon$. This interval is called the period. The reciprocal of the period is the frequency ν, and λ is the wavelength. It is convenient to let $K = 2\pi/\lambda$, so that $f(K\varnothing) = f(2\pi z/\lambda \pm 2\pi\nu t)$. Thus, υ and ν are related by

$$\upsilon = \nu\lambda. \tag{2.20}$$

A large class of problems can be solved by representing the waves by sinusoidally varying functions of the form $f = f_0 \sin 2\pi(z/\lambda \pm \nu t)$ or $f = f_0 \cos 2\pi(z/\lambda \pm \nu t)$. The reason sinusoidal solutions are so useful is that the principles of Fourier analysis and synthesis allow an arbitrary function to be described mathematically by sums of sinusoidal waves of appropriate amplitude and phase, provided the function is mathematically well behaved. In practice this qualification is not greatly restrictive, because virtually any function that describes a physically real quantity will be well behaved.

Because exponential functions are easy to manipulate mathematically, it is often convenient to use complex variables to represent the sinusoidal functions through the relation

$$e^{i\theta} = \cos\theta + i\sin\theta, \tag{2.21}$$

where $i = \sqrt{-1}$ and θ is some arbitrary quantity. Once the complex solution has been obtained, only the real part represents an actual physical quantity. Thus, solutions of the wave equation of the form $\mathbf{E}_e = \mathbf{E}_{e0}e^{2\pi i(z/\lambda - \nu t)}$ and $\mathbf{B}_m = \mathbf{B}_{m0}e^{2\pi i(z/\lambda - \nu t)}$ may be investigated without loss of generality. Comparing (2.14) and (2.15) with (2.16) shows that these expressions for the fields represent plane sinusoidal waves moving in the positive z direction with velocity

$$\upsilon = (\mu_{m0}\varepsilon_{e0})^{-1/2} = c_0, \tag{2.22}$$

where c_0 is the velocity of light in free space ($c_0 = 2.998 \times 10^8\,\mathrm{m\,s^{-1}}$). The quantities $\mathbf{E}_{e0}$ and $\mathbf{B}_{m0}$ are the amplitudes of the fields and are determined by the boundary conditions.

The electric and magnetic fields are not independent, but are related through (2.11), which for a plane sinusoidal wave propagating in the z direction is

$$\mathrm{curl}\,\mathbf{E}_e = \mathbf{u}_z \times \partial\mathbf{E}_e/\partial z = (2\pi i/\lambda)\mathbf{u}_z \times \mathbf{E}_e = -\partial\mathbf{B}_m/\partial t = 2\pi i v\mathbf{B}_m. \quad (2.23)$$

Thus,

$$\mathbf{B}_m = \mathbf{u}_z \times \mathbf{E}_e/c_0. \quad (2.24)$$

Similarly (2.12) is

$$\mathbf{u}_z \times \partial\mathbf{B}_m/\partial z = i\,(2\pi/\lambda)\,\mathbf{u}_z \times \mathbf{B}_m = \mu_{m0}\varepsilon_{e0}\partial\mathbf{E}_e/\partial t = -2\pi i v\mu_{m0}\varepsilon_{e0}\mathbf{E}_e, \quad (2.25)$$

so that

$$\mathbf{E}_e = -(1/\lambda v\mu_{m0}\varepsilon_{e0})\mathbf{u}_z \times \mathbf{B}_m = -c_0\mathbf{u}_z \times \mathbf{B}_m. \quad (2.26)$$

Equations (2.24) and (2.26) show that $\mathbf{E}_e$ and $\mathbf{B}_m$ are perpendicular to each other and to the direction of propagation (Figure 2.2), and also that the amplitudes of the fields are related by

$$B_{m0} = E_{e0}/c_0. \quad (2.27)$$

Two independent orthogonal solutions to the wave equation are possible in which the electric vectors are perpendicular to each other. If the positive x direction is chosen to be parallel to $\mathbf{E}_e$, then the component of $\mathbf{B}_m$ corresponding to this solution points in the positive y direction. Figure 2.2 illustrates this solution. If the positive y direction is chosen to be parallel to the electric vector, then the accompanying magnetic vector points in the negative x direction. In free space, neither $\mathbf{E}_e$ nor $\mathbf{B}_m$ has a component parallel to z.

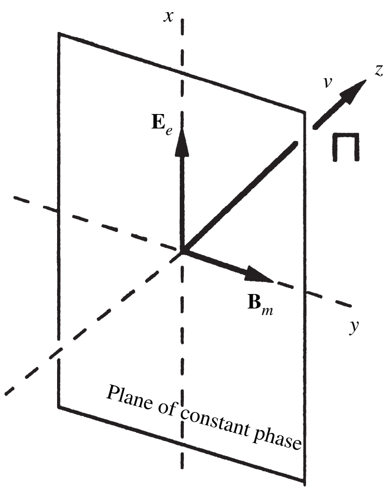

Figure 2.2 Relation between the fields and the propagation velocity vector in a plane electromagnetic wave.

2.2.2 Huygens's principle

Three hundred years ago when it was realized that light had a wave nature, the question arose as to what medium the waves were propagating through. It was postulated that all space was filled with an invisible fluid called the "aether." However, in the modern view of an electromagnetic wave no aether is required. According to Maxwell's equations, a changing electric field generates a magnetic field, and a changing magnetic field generates an electric field. Thus, a propagating electromagnetic wave may be regarded as generating itself; that is, the changing fields on the wave front continually regenerate the wave. This concept leads to *Huygens's principle*, in which each point on a wave front may be considered to be a source of spherical wavelets that travel radially outward and combine coherently with wavelets from all the other points to produce a new wave front. If the wave front is plane and infinite in lateral extent, this process simply produces another plane wave front. However, if part of the wave front is obstructed, Huygens's principle predicts that fields still exist behind the obstructing object. This phenomenon is called *diffraction*, and Huygen's principle can be used to calculate the resultant fields in the vicinity of the object.

2.2.3 The Poynting vector and the irradiance

An important quantity, the power contained in the wave, can be obtained by forming the vector dot products of **E** with equation (2.12) and of **B** with (2.13),

$$\mathbf{E}_e \cdot \operatorname{curl} \mathbf{B}_m = \mu_{m0}\varepsilon_e \mathbf{E}_e \cdot \partial \mathbf{E}_e/\partial t = (\mu_{m0}\varepsilon_{e0}/2)\partial \mathbf{E}_e^2/\partial t, \tag{2.28}$$

$$\mathbf{B}_m \cdot \operatorname{curl} \mathbf{E}_e = -\mathbf{B} \cdot \partial \mathbf{B}_m/\partial t = (1/2)\partial B_m^2/\partial t; \tag{2.29}$$

subtracting gives

$$\mathbf{E}_e \cdot \operatorname{curl} \mathbf{B}_m - \mathbf{B}_m \cdot \operatorname{curl} \mathbf{E}_e = \frac{1}{2}\partial(\mu_{m0}\varepsilon_{e0}E_e^2 + B_m^2)/\partial t. \tag{2.30}$$

Using the vector identity, equation (A.10), this becomes

$$(1/\mu_{m0})\operatorname{div}(\mathbf{E}_e \times \mathbf{B}_m) = -\frac{1}{2}\partial(\varepsilon_{e0}E_e^2 + B_m^2/\mu_{m0})/\partial t. \tag{2.31}$$

Integrating both sides of the last equation over a volume V bounded by a closed surface A (Figure 2.3), and applying Gauss's theorem [equation (A.11)], gives

$$\begin{aligned}\int_V \operatorname{div}(\mathbf{E}_e \times \mathbf{B}_m/\mu_{m0})dV &= \int_A (\mathbf{E}_e \times \mathbf{B}_m/\mu_{m0}) \cdot dA \\ &= -\frac{\partial}{\partial t}\int_V (\varepsilon_{e0}E_e^3/2 + B_m^2/2\mu_{m0})dV.\end{aligned} \tag{2.32}$$