Chapter 5 | Light Scattering by Atmospheric Particulates

5.1 Morphology of Atmospheric Particulates

The earth's atmosphere contains various types of particulates ranging from aerosols, water droplets, and ice crystals to raindrops, snowflakes, and hailstones. They are produced by a number of physical and dynamic processes, which govern their formation and growth in the atmosphere. In association with the discussion of the fundamentals of light scattering and absorption by particulates, we present in the following an overview of the morphology of pertinent atmospheric particulates, with a specific emphasis on aerosols and cloud particles.

In Section 3.1.2, we briefly discussed the sources of aerosol particles in the atmosphere. Aerosols of natural origin include dust from arid and semiarid regions, particles from sea spray over the oceans, volcanic debris, smoke from forest fires, extraterrestrial or interplanetary dust, and small particles produced by the chemical reactions of natural gases. Man-made aerosols are produced by particles directly emitted during combustion processes and particles formed from emitted gases.

Aerosols are usually classified in terms of their location and type. The continental type is subdivided into clear (rural, forest), average, desert (background, windy), and urban/industrial conditions. The aerosol components involved in this type are water-soluble particles, dustlike particles, soot, and minerals. The maritime type is grouped into clear, mineral, and polluted conditions, involving sea salt, mineral, sulfate, and soot particles. The polar type is subdivided into polluted, clear Arctic, and clear Antarctic conditions, which contain soot, mineral, sea salt, and sulfate particles. Because of the diversity of aerosol types in various regions, it appears that they must be governed in part by the transport process associated with atmospheric circulation.

The dustlike substances are mineral dust formed in nondesert locations and are representative of soil conditions. Water-soluble substances refer to the parts of aerosols that are soluble in water and consist of a mixture of sulfate, nitrate, and organic compounds, the mixing ratio of which varies with the source. The term "soot" represents all carbonaceous materials that are a product either of direct particle emission into the atmosphere by combustion processes or of the transformation of combustion-related gases to particles. Sulfate particles are products of the oxidation of sulfur-bearing gases

that are generated both naturally and anthropogenically. Natural sulfate is produced mostly by biogenic sources, whereas its anthropogenic counterpart stems primarily from coal and oil burning, smelting, petroleum refining, and transportation. The diverse tropospheric aerosol types have been conveniently grouped into five basic categories: dustlike soil, soot, sulfate, sea salt, and organic aerosols.

The size distributions of atmospheric aerosols are complex, but they have frequently been divided into two size classes, representing two primary formation mechanisms. Particles with diameters larger than about 1 μm are produced by the breakup and suspension of bulk materials by the wind (e.g., sea salt and soil dust). Fine particles smaller than 1 μm are usually formed by combustion or chemical conversions of the gaseous precursors into liquid or solid products. The size spectrum of aerosols in the atmosphere has been the subject of extensive research in the past four decades, and a detailed discussion is beyond the scope of this text. However, for the purpose of illustration, Fig. 5.1 depicts well-known measured size distributions of natural aerosols in reference to the logarithmic scale (Junge, 1963). Aerosols with radii smaller than 0.1 μm are collectively referred to as *Aitken nuclei.* For radii between 0.1 and 1.0 μm, the aerosols are referred to as *large nuclei.* Aerosols with radii larger than 1.0 μm are called *grant nuclei.* It is clear that the majority of atmospheric aerosols have sizes on the order of 0.1 μm, but aerosols as large as 10 to 20 μm have also been observed. To facilitate the remote sensing of aerosols, a number of analytic expressions have been developed to represent their size distributions. These include the Junge power law (Fig. 5.1 and Section 7.2.1), the gamma distribution [Eq. (5.3.44)], the log-normal distribution [Eq. (5.3.45)], and their modifications.

The radiative properties of atmospheric aerosols are critically dependent on their refractive indices as functions of wavelength. Both the real and imaginary parts of aerosols have been measured in the laboratory and tabulated in terms of dustlike, water-soluble, soot, oceanic, sulfate, mineral, and water substances covering a spectrum of wavelengths (e.g., d'Almeida *et al.,* 1991). In the solar visible region, absorption due to oceanic and sulfate particles is relatively small. A large absorption occurs for mineral, dustlike, and water-soluble particles, and particularly for soot. In a wet environment, aerosols interact with the ambient water vapor, the process of which affects their size, shape, and chemical composition, and consequently their optical properties. The humidity effects on aerosol size, shape, and composition are intricate and are directly related to the formation of water droplets.

Electron microscopic photographs of aerosols show a great variety of shapes ranging from quasi-spherical to highly irregular geometries. Solid aerosols are generally irregular. Figure 5.2 displays an example of microsized dust particles from the Sahara desert collected in Israel. The shape of aerosols also depends on the relative humidity. Many spherical aerosols have internal inclusions and/or attachments (see Fig. 5.28). A large number of submicrometer particles from combustion and biomass burnings are in the form of clusters and aggregates. Some aerosols are effective condensation and ice nuclei upon which cloud particles are formed in the atmosphere.
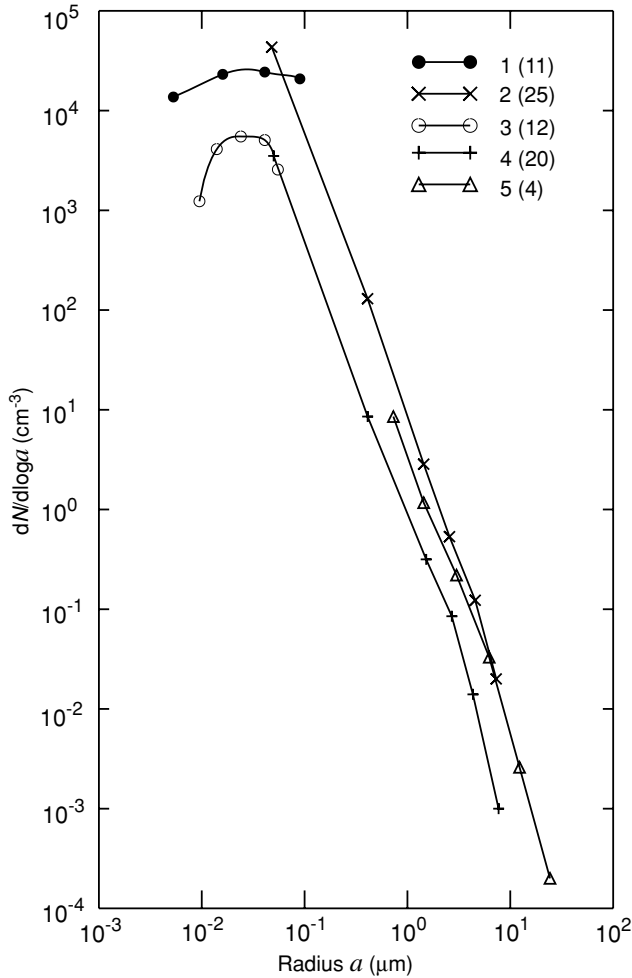
Figure 5.1 Size distribution of natural aerosols. Curves 1, 2, and 5 were measurements at Frankfurt/Main. Curve 1 represents ion counts converted to nuclei numbers. Curve 2 denotes data from impactors, and curve 5 is the average sedimentation data. Curves 3 and 4 were simultaneous measurements at Zugspitze, 3 km above sea level. The figures in parentheses are the number of individual measurements (data taken from Junge, 1963).

Clouds, which are composed of water droplets and/or ice crystals, are conventionally classified in terms of their position and appearance in the atmosphere. In midlatitudes, clouds with base heights of about ∼6 km are defined as high clouds and are commonly referred to as cirrus clouds. The group of low clouds with base heights below ∼2 km include stratus and cumulus. In between high and low clouds are a group called middle clouds consisting of altocumulus and altostratus. Clouds with significant vertical developments such as those occurring in the tropics are named cumulonimbus.
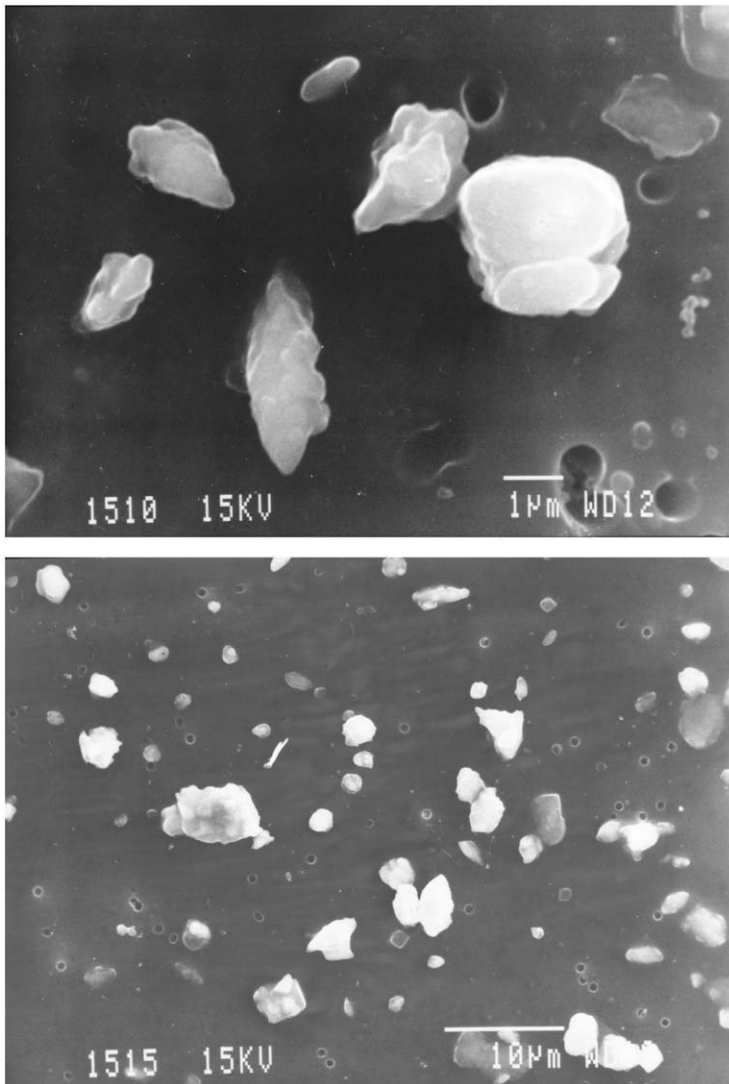
Figure 5.2   Microsized dust particles from the Sahara desert collected in Israel in 1998. Note that the scales for upper and lower diagrams are 1 and 10 μm, respectively (courtesy of Yoram Kaufman).

The microphysical state of a cloud is determined by the dispersion of particle sizes and their phase (ice or liquid). Based on extensive aircraft observations, low clouds and some middle clouds, because of temperature stratification, are generally composed of spherical water droplets with sizes ranging from $\sim$1 μm to 20 μm. A typical water droplet is on the order of 5 μm. Some middle clouds with temperatures warmer than about $-20°$C contain supercooled water droplets that coexist with ice particles.

Small water droplets are spherical in nature because of the requirement of surface pressure to hold the water molecules together. However, collision and coalescence of a fortunate larger water droplet (on the order of 20 μm) with other water droplets can produce millimeter- and centimeter-sized raindrops, as is evident from our daily experience. The falling raindrops in a dynamic flow field deviate from the spherical shape and have been investigated by Pruppacher and Pitter (1971) and Wang (1982). The latter author has developed mathematical expressions to define the shape of falling graupel, hailstones, and some deformed raindrops. The nonsphericity of these hydrometeors is important in the development of radar backscattering techniques for the quantitative detection of precipitation and thunderstorms based on light scattering theory.

Ice crystals in the atmosphere are primarily present in cirrus clouds, as well as in the top portion of some middle clouds. The formation, maintenance, and dissipation of cirrus clouds are principally associated with large-scale synoptic features and disturbances. In the tropics, they are related to deep-cumulus outflows. The ice crystal shapes depend on temperature and relative humidity as well as whether they undergo collision and coalescence processes in the clouds. Weickmann (1948) observed that at humidities approaching water saturation, ice crystals in cirrus castellatus, cirrocumulus, and cirrus generating cells have prismatic skeleton shapes that occur in hollow and cluster crystals, referred to as *bullet rosettes.* Between ice and water saturation, ice crystals grow in the form of prisms in cirrus filosus and cirrus densus. In cirrostratus, where relative humidities are close to ice saturation, ice crystals are primarily individual with full crystals such as columns, prisms, and plates.

In midlatitude cirrus, where substantial aircraft observations have been conducted, ice crystal shapes are generally classified as a function of temperature (Heymsfield and Platt, 1984; Heymsfield and Iaquinta, 2000). Hollow columns and hexagonal plates are the most abundant types near the cloud top for most cirrus clouds. Spatial ice crystals such as bullet rosettes are the predominant forms above about $-40°$C, while hollow or solid columns prevail below about $-50°$C. In between these temperatures, convective cirrus contain predominantly spatial crystal forms, whereas stable cirrus are primarily composed of hollow columns. Figure 5.3 illustrates a spectrum of ice crystal sizes and shapes as a function of height in a typical midlatitude cirrus. It is evident that at the cloud top, pristine and small columns and plates are predominant. At the cloud bottom, however, irregular bullet rosettes and aggregates are produced, presumably due to the collision and coalescence associated with vertical mixing and gravitational pulling. Because ice crystal size and shape vary greatly with time and space, a presentation of representative values for the purpose of remote sensing and climate applications is a difficult task. Nevertheless, typical ice-crystal size distributions for midlatitude cirrus clouds have been developed on the basis of a number of intensive field observations of cirrus clouds from aircraft platforms, including the First ISCCP Regional Experiments in 1986 and 1991, the European cirrus experiment in 1989, and the Subsonic Aircraft: Contrail and Cloud Effect Special Study in 1996. These distributions are shown in Fig. 5.4. Displayed are five representative
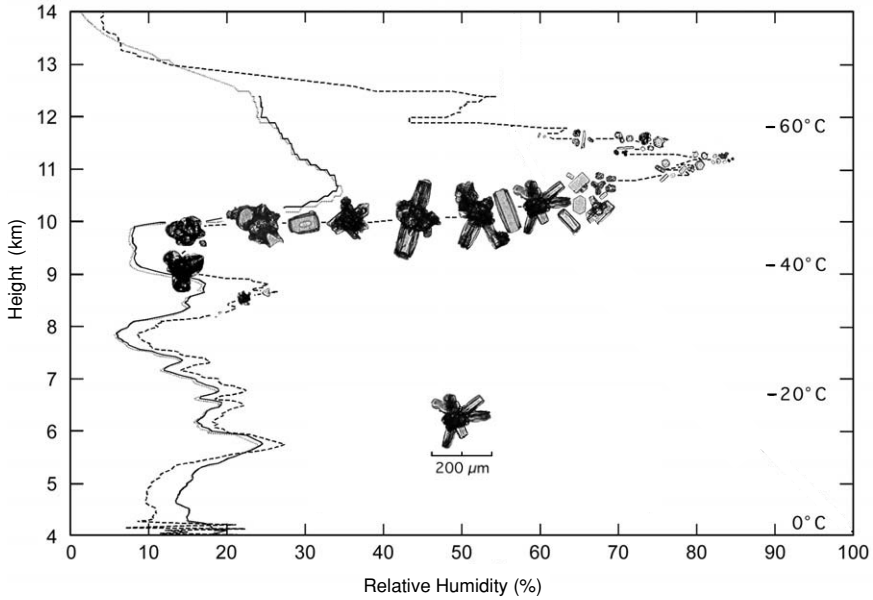
Figure 5.3   Ice crystal size and shape as a function of height and relative humidity captured by a replicator balloon sounding system in Marshall, Colorado, on November 10, 1994. The relative humidity was measured by cryogenic hygrometer (dashed line) and Vaisala RS80 instruments (solid line and dots). Also shown is temperature as a function of height (courtesy of Andrew Heymsfield of the National Center for Atmospheric Research).

size distributions spanning a mean effective ice crystal size range from 10 to 124 μm defined by

$$D_e = \int_{L_{\min}}^{L_{\max}} V\, n(L)\, dL \Big/ \int_{L_{\min}}^{L_{\max}} A\, n(L)\, dL, \qquad (5.1.1)$$

where V is the volume of an ice crystal, L is the maximum dimension, and A is the geometric projected area of an ice crystal on a surface perpendicular to the incident beam. For the purpose of calculating D_e, we have assumed $V \approx LD^2$ and $A \approx LD$, where D is the ice-crystal width. From available replicator and optical probe observations, the majority of ice-crystal shapes are aggregate, bullet rosettes, hollow columns, and plates, as displayed in Fig. 5.3.

Observations of ice-crystal size and shape distributions have been extremely limited for tropical cirrus. During the Central Equatorial Pacific Experiment conducted in 1993, aircraft microphysical measurements were carried out by a two-dimensional optical probe (30–300 μm) and a video ice particle sampler (<30 μm). Based on available measurements, ice-crystal sizes in the tropics range from about 10 μm to 2000 μm with four predominant shapes: bullet rosettes, aggregates, hollow columns, and plates, similar to those occurring in midlatitudes. Smaller ice-crystal sizes generally occur in cloud-top (colder temperature) conditions, whereas larger ice-crystal
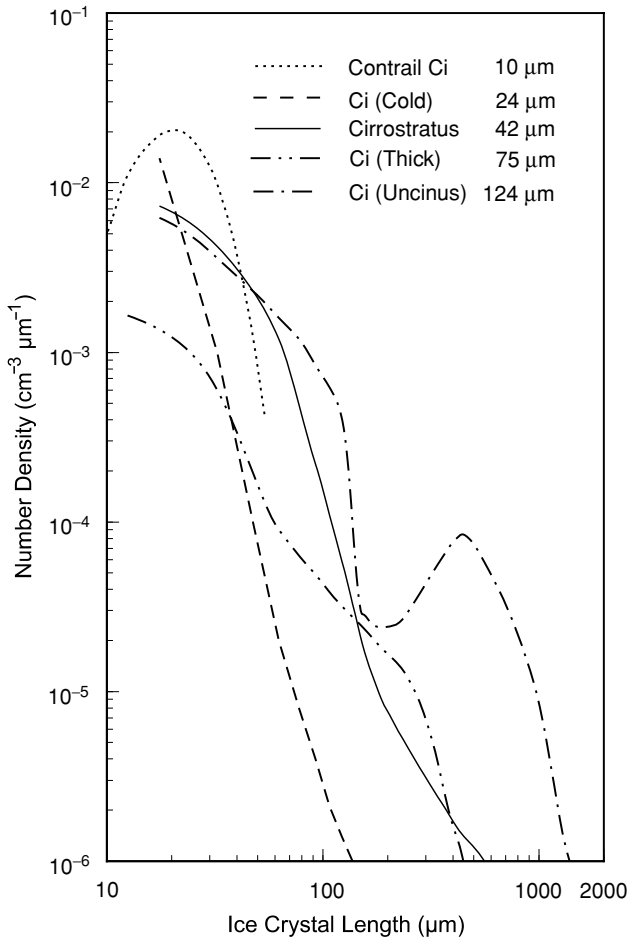
Figure 5.4   Representative ice-crystal size distributions for midlatitude cirrus clouds covering a range of mean effective ice-crystal sizes from 10 μm (Contrail), 24 μm (Cold), 42 μm (Cs), 75 μm (Thick), to 124 μm (Uncinus) (data from Heymsfield and Platt, 1984; First ISCCP Regional Experiment, 1991; Liou *et al.,* 1998).

sizes are associated with warmer temperatures and/or the developing stage of clouds associated with convection (Rolland *et al.,* 2000).

Ice-crystal shape and size data in arctic cirrus and stratiform clouds associated with frontal systems was collected during the First ISCCP Regional Experiment Arctic Cloud Experiment. Ice crystals with sizes larger than 40 μm were collected, and inspection of their shapes shows a combination of pristine and irregular types, including solid and hollow columns, prisms, plates, aggregates, and branched particles (Korolev *et al.,* 1999). The extensive collection of ice particles at a surface station in the Antarctic illustrated the prevalence of long needle ice-crystal types (Grenfell and

<div align="center">Table 5.1</div>

Sizes and Size Parameters for Atmospheric Particulates in Visible, Infrared, and Microwave Wavelengths

Type	Size (a)	Size parameters ($2\pi a/\lambda$)		
		$\lambda_v(0.5\,\mu m)$	$\lambda_{ir}(10\,\mu m)$	$\lambda_m(1\,cm)$
Aerosol (S*/NS†)	$\lesssim 1\,\mu m$	1.26×10^1	6.3×10^{-1}	6.3×10^{-4}
Water droplet (S)	$\sim 10\,\mu m$	1.26×10^2	6.3×10^0	6.3×10^{-3}
Ice crystal (NS)	$\sim 10^2\,\mu m$	1.26×10^3	6.3×10^1	6.3×10^{-2}
Raindrop (NS)	$\sim 1\,mm$	1.26×10^4	6.3×10^2	6.3×10^{-1}
Snowflake (hailstone) (NS)	$\sim 1\,cm$	1.26×10^5	6.3×10^3	6.3×10^0

*Spherical; †nonspherical.

Warren, 1999). Clearly, then, ice crystals vary substantially in size and shape from the tropics to midlatitudes to the polar regions.

In Table 5.1, we summarize typical sizes of atmospheric particulates in terms of an equivalent radius (for nonspherical particles) and size parameters (Section 1.1.4) with respect to representative wavelengths in the solar, infrared, and microwave regions. Except in the microwave region, all the particle sizes are comparable to or larger than solar and infrared wavelengths. In these cases, the dipole mode of the electric field, which leads to the development of the Rayleigh scattering theory (Section 3.3.1), is not applicable.

In Section 3.3.2, we briefly introduced the Lorenz–Mie theory, the geometric optics approach, and the anomalous diffraction theory of light scattering by particles. This chapter is a continuation of the discussion of atmospheric scattering. We first present a complete description of the Lorenz–Mie theory of light scattering by spheres, which is applicable to spherical aerosols and cloud droplets. Next, we introduce the fundamentals of the geometric optics approach in terms of diffraction and geometric reflection and refraction. Finally, we present contemporary developments in the study of light scattering by nonspherical ice crystals and aerosols.

5.2   Lorenz–Mie Theory of Light Scattering by Spherical Particles

5.2.1   Electromagnetic Wave Equation and Solution

We shall first introduce the fundamental Maxwell equations for the electromagnetic field. The state of excitation that is established in space by the presence of electric charges is said to constitute an electromagnetic field. It is represented by two vectors **E** and **B,** called the electric vector and magnetic induction, respectively. It is necessary to introduce a second set of vectors involving the electric current density **j,** the electric displacement **D,** and the magnetic vector **H** to describe the effects of the electromagnetic field on material objects. At every point where the physical properties of the medium are continuous in its neighborhood, the space and time derivatives of these

five vectors can be related by Maxwell's equations as follows:

$$\nabla \times \mathbf{H} = \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} + \frac{4\pi}{c}\mathbf{j}, \tag{5.2.1}$$

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \tag{5.2.2}$$

$$\nabla \cdot \mathbf{D} = 4\pi\rho, \tag{5.2.3}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{5.2.4}$$

where t denotes time, c the velocity of light, and ρ the density of charge. Equation (5.2.3) may be regarded as a defining equation for the electric charge density ρ, and Eq. (5.2.4) implies that no free magnetic poles exist.

From Eq. (5.2.1), since $\nabla \cdot \nabla \times \mathbf{H} = 0$, a dot-product operation leads to

$$\nabla \cdot \mathbf{j} = -\frac{1}{4\pi}\nabla \cdot \frac{\partial \mathbf{D}}{\partial t}. \tag{5.2.5}$$

Hence, differentiating Eq. (5.2.3) with respect to t leads to

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{5.2.6}$$

This is the equation of continuity in an electromagnetic field.

To allow a unique determination of the field vectors from a given distribution of current and charges, the preceding equations must be supplemented by relationships describing the behavior of substances under the influence of the field. These relationships are given by

$$\mathbf{j} = \sigma \mathbf{E}, \tag{5.2.7}$$

$$\mathbf{D} = \varepsilon \mathbf{E}, \tag{5.2.8}$$

$$\mathbf{B} = \mu \mathbf{H}, \tag{5.2.9}$$

where σ is the specific conductivity, ε the permittivity, and μ the magnetic permeability.

We shall now confine our attention to the field where there are no charges ($\rho = 0$) and current ($|\mathbf{j}| = 0$), and to the medium which is homogeneous so that ε and μ are constants. Thus, the Maxwell equations reduce to

$$\nabla \times \mathbf{H} = \frac{\varepsilon}{c}\frac{\partial \mathbf{E}}{\partial t}, \tag{5.2.10}$$

$$\nabla \times \mathbf{E} = -\frac{\mu}{c}\frac{\partial \mathbf{H}}{\partial t}, \tag{5.2.11}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{5.2.12}$$

$$\nabla \cdot \mathbf{H} = 0. \tag{5.2.13}$$

Equations (5.2.10)–(5.2.13) will be used in the following to derive the electromagnetic wave equation. Note that Eqs. (5.2.12) and (5.2.13) can be obtained directly from Eqs. (5.2.10) and (5.2.11) by carrying out the dot operation.

We shall consider a plane electromagnetic wave in a periodic field with a circular frequency ω so that we may write

$$\mathbf{E} \rightarrow \mathbf{E}e^{i\omega t}, \tag{5.2.14}$$

$$\mathbf{H} \rightarrow \mathbf{H}e^{i\omega t}. \tag{5.2.15}$$

On the basis of these transformations, Eqs. (5.2.10) and (5.2.11) become

$$\nabla \times \mathbf{H} = ikm^2\mathbf{E}, \tag{5.2.16}$$

$$\nabla \times \mathbf{E} = -ik\mathbf{H}, \tag{5.2.17}$$

where $k = 2\pi/\lambda (= \omega/c)$ is the wavenumber denoting the propagation constant in vacuum, λ is the wavelength in vacuum, $m = \sqrt{\varepsilon}$ is the complex refractive index of the medium at the frequency ω, and $\mu \approx 1$ is the permeability of air. Note that the wavenumber k used in this chapter differs from the wavenumber ν (cm^{-1}) defined in Section 1.1 by a factor of 2π.

We then perform the curl operation of Eq. (5.2.17) to obtain

$$\nabla \times \nabla \times \mathbf{E} = -ik\nabla \times \mathbf{H}. \tag{5.2.18}$$

Moreover, by noting that $\nabla \cdot \nabla \times \mathbf{E} = 0$ and that $\nabla \cdot \mathbf{E} = 0$, we get

$$\nabla^2\mathbf{E} = -k^2m^2\mathbf{E}. \tag{5.2.19}$$

In a similar way, from Eqs. (5.2.16) and (5.2.13), we derive the following relationship:

$$\nabla^2\mathbf{H} = -k^2m^2\mathbf{H}. \tag{5.2.20}$$

Equations (5.2.19) and (5.2.20) reveal that the electric vector and magnetic induction in a homogeneous medium satisfy the vector wave equation in the form

$$\nabla^2\mathbf{A} + k^2m^2\mathbf{A} = 0, \tag{5.2.21}$$

where $\mathbf{A}$ can be either $\mathbf{E}$ or $\mathbf{H}$.

Now, if ψ satisfies the scalar wave equation

$$\nabla^2\psi + k^2m^2\psi = 0, \tag{5.2.22}$$

vectors $\mathbf{M}_\psi$ and $\mathbf{N}_\psi$ in spherical coordinates (r, θ, ϕ) defined by

$$\mathbf{M}_\psi = \nabla \times [\mathbf{a}_r(r\psi)] = \left(\mathbf{a}_r\frac{\partial}{\partial r} + \mathbf{a}_\theta\frac{1}{r}\frac{\partial}{\partial\theta} + \mathbf{a}_\phi\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right) \times [\mathbf{a}_r(r\psi)]$$

$$= \mathbf{a}_\theta\frac{1}{r\sin\theta}\frac{\partial(r\psi)}{\partial\phi} - \mathbf{a}_\phi\frac{1}{r}\frac{\partial(r\psi)}{\partial\theta}, \tag{5.2.23}$$

$$mk\mathbf{N}_\psi = \nabla \times \mathbf{M}_\psi$$

$$= \mathbf{a}_r r\left[\frac{\partial^2(r\psi)}{\partial r^2} + m^2k^2(r\psi)\right] + \mathbf{a}_\theta\frac{1}{r}\frac{\partial^2(r\psi)}{\partial r\partial\theta} + \mathbf{a}_\phi\frac{1}{r\sin\theta}\frac{\partial^2(r\psi)}{\partial r\partial\phi}$$

$$\tag{5.2.24}$$

satisfy the vector wave equation defined in Eq. (5.2.21) subject to Eq. (5.2.22). The terms $\mathbf{a}_r$, $\mathbf{a}_\theta$, and $\mathbf{a}_\phi$ are unit vectors in spherical coordinates. To obtain Eq. (5.2.24), we used Eq. (5.2.29), which is defined in the following.

Assuming that u and v are two independent solutions of the scalar wave equation defined in Eq. (5.2.22), then the electric and magnetic field vectors expressed by

$$\mathbf{E} = \mathbf{M}_v + i\mathbf{N}_u, \tag{5.2.25}$$

$$\mathbf{H} = m(-\mathbf{M}_u + i\mathbf{N}_v) \tag{5.2.26}$$

satisfy Eqs. (5.2.16) and (5.2.17). Employing Eqs. (5.2.23) and (5.2.24), $\mathbf{E}$ and $\mathbf{H}$ can be written explicitly as follows:

$$\mathbf{E} = \mathbf{a}_r \frac{i}{mk}\left[\frac{\partial^2(ru)}{\partial r^2} + m^2 k^2(ru)\right] + \mathbf{a}_\theta\left[\frac{1}{r\sin\theta}\frac{\partial(rv)}{\partial\phi} + \frac{i}{mkr}\frac{\partial^2(ru)}{\partial r\,\partial\theta}\right]$$
$$+ \mathbf{a}_\phi\left[-\frac{1}{r}\frac{\partial(rv)}{\partial\theta} + \frac{1}{mkr\sin\theta}\frac{\partial^2(ru)}{\partial r\,\partial\phi}\right], \tag{5.2.27}$$

$$\mathbf{H} = \mathbf{a}_r \frac{i}{k}\left[\frac{\partial^2(rv)}{\partial r^2} + m^2 k^2(rv)\right] + \mathbf{a}_\theta\left[-\frac{m}{r\sin\theta}\frac{\partial(ru)}{\partial\phi} + \frac{i}{kr}\frac{\partial^2(rv)}{\partial r\,\partial\theta}\right]$$
$$+ \mathbf{a}_\phi\left[\frac{m}{r}\frac{\partial(ru)}{\partial\theta} + \frac{i}{kr\sin\theta}\frac{\partial^2(rv)}{\partial r\,\partial\phi}\right]. \tag{5.2.28}$$

The scalar wave equation defined in Eq. (5.2.22) may be written in spherical coordinates in the form

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{r^2\sin\theta}\frac{\partial^2\psi}{\partial\phi^2} + k^2 m^2\psi = 0. \tag{5.2.29}$$

This equation is separable by defining

$$\psi(r,\theta,\phi) = R(r)\Theta(\theta)\Phi(\phi). \tag{5.2.30}$$

Upon substituting Eq. (5.2.30) into Eq. (5.2.29) and dividing the entire equation by $\psi(r,\theta,\phi)$, we obtain

$$\frac{1}{r^2}\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{1}{\Theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)$$
$$+ \frac{1}{r^2\sin^2\theta}\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2} + k^2 m^2 = 0. \tag{5.2.31}$$

We then multiply Eq. (5.2.31) by $r^2\sin^2\theta$ to obtain

$$\left[\sin^2\theta\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + \sin\theta\frac{1}{\Theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)\right.$$
$$\left. + k^2 m^2 r^2\sin^2\theta\right] + \frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2} = 0. \tag{5.2.32}$$

Since the first three terms in this equation consist of the variables r and θ, but not ϕ, the only possibility that Eq. (5.2.32) can be valid is when

$$\frac{1}{\Phi}\frac{d^2\Phi}{d\phi^2} = \text{const} = -\ell^2, \tag{5.2.33}$$

where we set the constant equal to $-\ell^2$ (ℓ denotes an integer) for mathematical convenience. In view of Eqs. (5.2.32) and (5.2.33), it is also clear that

$$\sin^2\theta\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + \sin\theta\frac{1}{\Theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right) + k^2m^2r^2\sin^2\theta - \ell^2 = 0. \tag{5.2.34}$$

On dividing Eq. (5.2.34) by $\sin^2\theta$, we obtain

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + k^2m^2r^2 + \frac{1}{\sin\theta}\frac{1}{\Theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right) - \frac{\ell^2}{\sin^2\theta} = 0. \tag{5.2.35}$$

Thus, we must have

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + k^2m^2r^2 = \text{const} = n(n+1), \tag{5.2.36}$$

$$\frac{1}{\sin\theta}\frac{1}{\Theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - \frac{\ell^2}{\sin^2\theta} = \text{const} = -n(n+1) \tag{5.2.37}$$

in order to satisfy Eq. (5.2.35), where n is an integer. The selection of the constant here is also for mathematical convenience. Rearranging Eqs. (5.2.33), (5.2.36), and (5.2.37) leads to

$$\frac{d^2(rR)}{dr^2} + \left[k^2m^2 - \frac{n(n+1)}{r^2}\right](rR) = 0, \tag{5.2.38}$$

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) + \left[n(n+1) - \frac{\ell^2}{\sin^2\theta}\right]\Theta = 0, \tag{5.2.39}$$

$$\frac{d^2\Phi}{d\phi^2} + \ell^2\Phi = 0. \tag{5.2.40}$$

The single-value solution for Eq. (5.2.40) is given by

$$\Phi = a_\ell\cos\ell\phi + b_\ell\sin\ell\phi, \tag{5.2.41}$$

where a_ℓ and b_ℓ are arbitrary constants. Equation (5.2.39) is the well-known equation for spherical harmonics. For convenience we introduce a new variable $\mu = \cos\theta$ so that

$$\frac{d}{d\mu}\left[(1-\mu^2)\frac{d\Theta}{d\mu}\right] + \left[n(n+1) - \frac{\ell^2}{1-\mu^2}\right]\Theta = 0. \tag{5.2.42}$$

The solutions of Eq. (5.2.42) can be expressed by the associated Legendre polynomials (spherical harmonics of the first kind) in the form

$$\Theta = P_n^\ell(\mu) = P_n^\ell(\cos\theta). \tag{5.2.43}$$

Finally, in order to solve the remaining equation (5.2.38), we set

$$kmr = \rho, \qquad R = (1/\sqrt{\rho})Z(\rho), \tag{5.2.44}$$

to obtain

$$\frac{d^2 Z}{d\rho^2} + \frac{1}{\rho}\frac{dZ}{d\rho} + \left[1 - \frac{(n + \frac{1}{2})^2}{\rho^2}\right] Z = 0. \tag{5.2.45}$$

The solution of this equation can be expressed by the general cylindrical function of order $n + \frac{1}{2}$ and is given by

$$Z = Z_{n+\frac{1}{2}}(\rho). \tag{5.2.46}$$

Thus, the solution of Eq. (5.2.38) is then

$$R = \frac{1}{\sqrt{kmr}} Z_{n+\frac{1}{2}}(kmr). \tag{5.2.47}$$

Upon combining Eqs. (5.2.41), (5.2.43), and (5.2.47), the elementary wave functions at all points on the surface of a sphere can be expressed by

$$\psi(r, \theta, \phi) = \frac{1}{\sqrt{kmr}} Z_{n+\frac{1}{2}}(kmr) P_n^{\ell}(\cos\theta)(a_{\ell}\cos\ell\phi + b_{\ell}\sin\ell\phi). \tag{5.2.48}$$

Each cylindrical function denoted in Eq. (5.2.47) may be expressed as a linear combination of two cylindrical functions of standard type, i.e., the Bessel function $J_{n+\frac{1}{2}}$ and the Neumann function $N_{n+\frac{1}{2}}$. We define

$$\psi_n(\rho) = \sqrt{\pi\rho/2}\,J_{n+\frac{1}{2}}(\rho), \qquad \chi_n(\rho) = -\sqrt{\pi\rho/2}\,N_{n+\frac{1}{2}}(\rho). \tag{5.2.49}$$

The functions ψ_n are regular in every finite domain of the ρ plane including the origin, whereas the functions χ_n have singularities at the origin $\rho = 0$ at which they become infinite. Thus, we may use ψ_n, but not χ_n, to represent the wave inside the sphere. On utilizing the definitions in Eq. (5.2.49), Eq. (5.2.47) can be rewritten in the form

$$rR = c_n \psi_n(kmr) + d_n \chi_n(kmr), \tag{5.2.50}$$

where c_n and d_n are arbitrary constants. Equation (5.2.50) now represents the general solution of Eq. (5.2.38). It follows that the general solution of the scalar wave equation (5.2.29) can then be expressed by

$$r\psi(r, \theta, \phi) = \sum_{n=0}^{\infty}\sum_{\ell=-n}^{n} P_n^{\ell}(\cos\theta)[c_n\,\psi_n(kmr)$$

$$+ d_n\,\chi_n(kmr)](a_{\ell}\cos\ell\phi + b_{\ell}\sin\ell\phi). \tag{5.2.51}$$

Note that the electric and magnetic field vectors of electromagnetic waves can be subsequently derived from Eqs. (5.2.27) and (5.2.28).

Moreover, when $c_n = 1$ and $d_n = i$, we have

$$\psi_n(\rho) + i\chi_n(\rho) = \sqrt{\pi\rho/2}H_{n+\frac{1}{2}}^{(2)}(\rho) = \xi_n(\rho), \qquad (5.2.52)$$

where $H_{n+\frac{1}{2}}^{(2)}$ is the half-integral-order Hankel function of the second kind, which has the property of vanishing at infinity in the complex plane. This function is therefore suitable for the representation of the scattered wave.

5.2.2   Formal Scattering Solution

Having solved the vector wave equation, we can now discuss the scattering of a plane wave by a homogeneous sphere. For simplicity, we shall assume that outside the medium is vacuum ($m = 1$), that the material of the sphere has an index of refraction m, and that the incident radiation is linearly polarized. We select the origin of a rectangular system of coordinates at the center of the sphere, with the positive z axis along the direction of propagation of the incident wave. If the amplitude of the incident wave is normalized to unity, the incident electric and magnetic field vectors are given by

$$\mathbf{E}^i = \mathbf{a}_x e^{-ikz}, \qquad \mathbf{H}^i = \mathbf{a}_y e^{-ikz}, \qquad (5.2.53)$$

where $\mathbf{a}_x$ and $\mathbf{a}_y$ are unit vectors along the x and y axes, respectively.

The components of any vector, say $\mathbf{a}$, in the Cartesian system (x,y,z) may be transformed to the spherical polar coordinates (r,θ,ϕ) in the forms

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta. \qquad (5.2.54)$$

Based on the geometric relationship shown in Fig. 5.5, we find

$$\mathbf{a}_r = \mathbf{a}_x \sin\theta\cos\phi + \mathbf{a}_y \sin\theta\sin\phi + \mathbf{a}_z \cos\theta,$$

$$\mathbf{a}_\theta = \mathbf{a}_x \cos\theta\cos\phi + \mathbf{a}_y \cos\theta\sin\phi - \mathbf{a}_z \sin\theta,$$

$$\mathbf{a}_\phi = -\mathbf{a}_x \sin\phi + \mathbf{a}_y \cos\phi, \qquad (5.2.55)$$

where $\mathbf{a}_x$, $\mathbf{a}_y$, and $\mathbf{a}_z$ are unit vectors in Cartesian coordinates, while $\mathbf{a}_r$, $\mathbf{a}_\theta$, and $\mathbf{a}_\phi$ are unit vectors in spherical coordinates.

In view of the preceding analysis, the electric and magnetic field vectors of the incident wave may be written in the forms

$$E_r^i = e^{-ikr\cos\theta}\sin\theta\cos\phi,$$

$$E_\theta^i = e^{-ikr\cos\theta}\cos\theta\cos\phi,$$

$$E_\phi^i = -e^{-ikr\cos\theta}\sin\phi. \qquad (5.2.56)$$

$$H_r^i = e^{-ikr\cos\theta}\sin\theta\sin\phi,$$

$$H_\theta^i = e^{-ikr\cos\theta}\cos\theta\sin\phi,$$
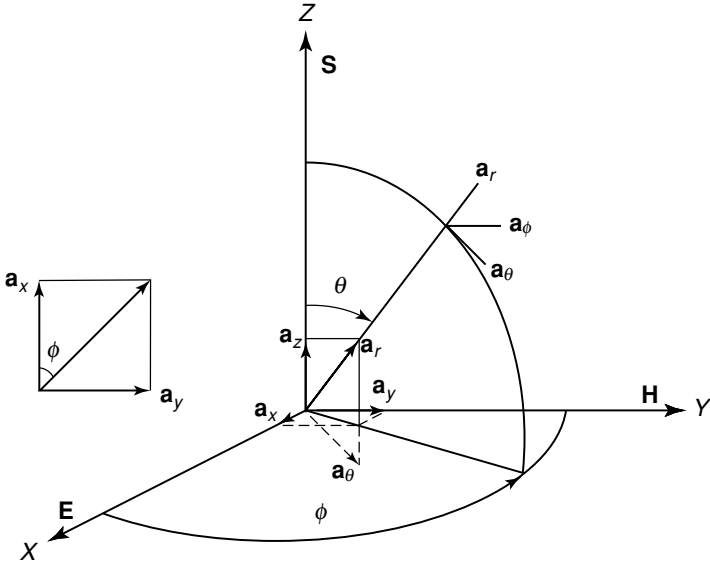
$$H_\phi^i = e^{-ikr\cos\theta}\cos\phi. \qquad (5.2.57)$$

Figure 5.5   Transformation of rectangular to spherical coordinates. **S** is the Poynting vector; **a** is an arbitrary unit vector; θ and ϕ are zenith and azimuthal angles; and **E** and **H** are electric and magnetic vectors placed on the X and Y axes, respectively.

On the basis of Bauer's formula (Watson, 1944), the first factor on the right-hand side of this equation may be expressed in the following differentiable series of Legendre polynomials:

$$e^{-ikr\cos\theta} = \sum_{n=0}^{\infty} (-i)^n (2n+1) \frac{\psi_n(kr)}{kr} P_n(\cos\theta), \tag{5.2.58}$$

where ψ_n is defined in Eq. (5.2.49). Moreover, we have the following mathematical identities:

$$e^{-ikr\cos\theta} \sin\theta = \frac{1}{ikr} \frac{\partial}{\partial\theta} (e^{-ikr\cos\theta}), \tag{5.2.59}$$

$$\frac{\partial}{\partial\theta} P_n(\cos\theta) = -P_n^1(\cos\theta), \qquad P_0^1(\cos\theta) = 0. \tag{5.2.60}$$

Equation (5.2.60) relates the Legendre polynomial P_n to the associated Legendre polynomial P_n^1.

To determine the potentials u and v, we require only one of the components in Eq. (5.2.27). The first of them is ($m = 1$)

$$E_r^i = e^{-ikr\cos\theta} \sin\theta\cos\phi = \frac{i}{k} \left[\frac{\partial^2(ru^i)}{\partial r^2} + k^2(ru^i) \right]. \tag{5.2.61}$$

In view of Eqs. (5.2.58)–(5.2.60), we have

$$e^{-ikr\cos\theta}\sin\theta\cos\phi = \frac{1}{(kr)^2}\sum_{n=1}^{\infty}(-i)^{n-1}(2n+1)\psi_n(kr)P_n^1(\cos\theta)\cos\phi.$$

(5.2.62)

Accordingly, we take a trial solution in Eq. (5.2.61) by using an expanding series of a similar form:

$$ru^i = \frac{1}{k}\sum_{n=1}^{\infty}\alpha_n\psi_n(kr)P_n^1(\cos\theta)\cos\phi.$$

(5.2.63)

Upon substituting Eqs. (5.2.62) and (5.2.63) into Eq. (5.2.61) and comparing coefficients, we obtain

$$\alpha_n\left[k^2\psi_n(kr) + \frac{\partial^2\psi_n(kr)}{\partial r^2}\right] = (-i)^n(2n+1)\frac{\psi_n(kr)}{r^2}.$$

(5.2.64)

In Eq. (5.2.50), since $\chi_n(kr)$ become infinite at the origin through which the incident wave must pass, we may let $c_n = 1$ and $d_n = 0$. It follows that

$$\psi_n(kr) = rR$$

(5.2.65)

is a solution of Eq. (5.2.38) (with $m = 1$),

$$\frac{d^2\psi_n}{dr^2} + \left(k^2 - \frac{\alpha}{r^2}\right)\psi_n = 0,$$

(5.2.66)

provided that $\alpha = n(n+1)$. Comparing Eq. (5.2.66) to (5.2.64), we find

$$\alpha_n = (-i)^n\frac{2n+1}{n(n+1)}.$$

(5.2.67)

Utilizing similar procedures, v^i can be derived from Eq. (5.2.28). Thus, for incident waves outside the sphere, we have

$$ru^i = \frac{1}{k}\sum_{n=1}^{\infty}(-i)^n\frac{2n+1}{n(n+1)}\psi_n(kr)P_n^1(\cos\theta)\cos\phi,$$

$$rv^i = \frac{1}{k}\sum_{n=1}^{\infty}(-i)^n\frac{2n+1}{n(n+1)}\psi_n(kr)P_n^1(\cos\theta)\sin\phi.$$

(5.2.68)

In order to match u^i and v^i with those of the internal and scattered waves whose potentials have been derived in Eq. (5.2.51), the latter must be expressed in a series of a similar form but with arbitrary coefficients. For internal waves, because the function $\chi_n(kmr)$ becomes infinite at the origin, only the function $\psi_n(kmr)$ may be used. Thus, for internal waves we have

$$ru^t = \frac{1}{mk}\sum_{n=1}^{\infty}(-i)^n\frac{2n+1}{n(n+1)}c_n\psi_n(kmr)P_n^1(\cos\theta)\cos\phi,$$

$$rv^t = \frac{1}{mk}\sum_{n=1}^{\infty}(-i)^n\frac{2n+1}{n(n+1)}d_n\psi_n(kmr)P_n^1(\cos\theta)\sin\phi.$$

(5.2.69)

For scattered waves, these two functions must vanish at infinity. The Hankel functions expressed in Eq. (5.2.52) possess precisely this property. Thus, for scattered waves we have

$$ru^s = \frac{1}{k} \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} a_n \xi_n(kr) P_n^1(\cos\theta) \cos\phi,$$

$$rv^s = \frac{1}{k} \sum_{n=1}^{\infty} (-i)^n \frac{2n+1}{n(n+1)} b_n \xi_n(kr) P_n^1(\cos\theta) \sin\phi. \qquad (5.2.70)$$

The coefficients a_n, b_n, c_n, and d_n must be determined from the boundary conditions at the surface of the sphere; that is, the tangential components of **E** and **H** are continuous across the spherical surface $r = a$ such that

$$E_\theta^i + E_\theta^s = E_\theta^t, \qquad H_\theta^i + H_\theta^s = H_\theta^t,$$

$$E_\phi^i + E_\phi^s = E_\phi^t, \qquad H_\phi^i + H_\phi^s = H_\phi^t. \qquad (5.2.71)$$

In view of Eqs. (5.2.27), (5.2.28), and (5.2.68)–(5.2.70), apart from common factors and differentiations with respect to θ and ϕ, which are the same for the wave inside and outside the sphere, both of the field components E_θ and E_ϕ contain the expressions v and $\partial(ru)/m\partial r$. It is also clear that the components H_θ and H_ϕ contain mu and $\partial(rv)/\partial r$. Equation (5.2.71) implies that these four expressions must be continuous at $r = a$. Consequently,

$$\frac{\partial}{\partial r}\left[r(u^i + u^s)\right] = \frac{1}{m}\frac{\partial}{\partial r}(ru^t), \qquad u^i + u^s = mu^t,$$

$$\frac{\partial}{\partial r}\left[r(v^i + v^s)\right] = \frac{\partial}{\partial r}(rv^t), \qquad v^i + v^s = v^t. \qquad (5.2.72)$$

From Eq. (5.2.72) we have

$$m[\psi_n'(ka) - a_n \xi_n'(ka)] = c_n \psi_n'(kma),$$

$$[\psi_n'(ka) - b_n \xi_n'(ka)] = d_n \psi_n'(kma),$$

$$[\psi_n(ka) - a_n \xi_n(ka)] = c_n \psi_n(kma),$$

$$m[\psi_n(ka) - b_n \xi_n(ka)] = d_n \psi_n(kma), \qquad (5.2.73)$$

where the prime denotes differentiation with respect to the argument. On eliminating c_n and d_n, we obtain the coefficients for the scattered waves in the forms

$$a_n = \frac{\psi_n'(y)\psi_n(x) - m\psi_n(y)\psi_n'(x)}{\psi_n'(y)\xi_n(x) - m\psi_n(y)\xi_n'(x)},$$

$$b_n = \frac{m\psi_n'(y)\psi_n(x) - \psi_n(y)\psi_n'(x)}{m\psi_n'(y)\xi_n(x) - \psi_n(y)\xi_n'(x)}, \qquad (5.2.74)$$

where $x = ka$, and $y = mx$. The solutions for c_n and d_n contain the same respective denominators as those of a_n and b_n but with $m[\psi_n'(x)\xi_n(x) - \psi_n(x)\xi_n'(x)]$ as a common numerator. At this point, solution of the scattering of electromagnetic waves by a

sphere with a radius of a and an index of refraction of m is complete. The electric and magnetic field vectors expressed in Eqs. (5.2.27) and (5.2.28) at any point inside or outside the sphere are now expressed in terms of the known mathematical functions given by Eqs. (5.2.68)–(5.2.70). For the sake of simplicity, we have assumed up to this point that the suspending medium is a vacuum. We may now let the outside medium and the sphere have refractive indices m_2 (real part) and m_1 (maybe complex), respectively. By replacing m by m_1/m_2 and the wavenumber k by $m_2 k$ (vacuum), the results given in Eq. (5.2.74) can be generalized to cases where a sphere is suspended in a medium.

5.2.3   The Far-Field Solution and Extinction Parameters

In the following, consider the scattered field at very large distances from the sphere. It is noted that for practical applications, light scattering observations are normally made in the far-field zone. In the far field, the Hankel functions denoted in Eq. (5.2.52) reduce to the form

$$\xi_n(kr) \approx i^{n+1} e^{-ikr}, \qquad kr \gg 1. \tag{5.2.75}$$

With this simplification, Eq. (5.2.70) becomes

$$ru^s \approx -\frac{ie^{-ikr}\cos\phi}{k} \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} a_n P_n^1(\cos\theta),$$

$$rv^s \approx -\frac{ie^{-ikr}\sin\phi}{k} \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} b_n P_n^1(\cos\theta). \tag{5.2.76}$$

The three components of the electric and magnetic field vectors in Eqs. (5.2.27) and (5.2.28) are given by

$$E_r^s = H_r^s \approx 0,$$

$$E_\theta^s = H_\phi^s \approx \frac{-i}{kr} e^{-ikr} \cos\phi \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left[a_n \frac{dP_n^1(\cos\theta)}{d\theta} + b_n \frac{P_n^1(\cos\theta)}{\sin\theta} \right],$$

$$E_\phi^s = H_\theta^s \approx \frac{i}{kr} e^{-ikr} \sin\phi \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left[a_n \frac{P_n^1(\cos\theta)}{\sin\theta} + b_n \frac{dP_n^1(\cos\theta)}{d\theta} \right].$$

$$\tag{5.2.77}$$

We find that the radial components E_r^s and H_r^s may be neglected in the far-field zone. To simplify Eq. (5.2.77), we define two *scattering functions* in the forms

$$S_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} [a_n \pi_n(\cos\theta) + b_n \tau_n(\cos\theta)],$$

$$S_2(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} [b_n \pi_n(\cos\theta) + a_n \tau_n(\cos\theta)], \tag{5.2.78}$$

where

$$\pi_n(\cos\theta) = \frac{1}{\sin\theta} P_n^1(\cos\theta),$$

$$\tau_n(\cos\theta) = \frac{d}{d\theta} P_n^1(\cos\theta). \tag{5.2.79}$$

Thus, we may write

$$E_\theta^s = \frac{i}{kr} e^{-ikr} \cos\phi \, S_2(\theta),$$

$$-E_\phi^s = \frac{i}{kr} e^{-ikr} \sin\phi \, S_1(\theta). \tag{5.2.80}$$

These fields represent an outgoing spherical wave where amplitude and the state of polarization are functions of the scattering angle θ. It is convenient to define the *perpendicular* and *parallel* components of the electric field as E_r and E_l, respectively. In reference to Fig. 5.6, the scattered perpendicular and parallel electric fields are given by

$$E_r^s = -E_\phi^s, \qquad E_l^s = E_\theta^s. \tag{5.2.81}$$

Also, the normalized incident electric vector [see Eq. (5.2.53)] may be decomposed into perpendicular and parallel components as follows:

$$E_r^i = e^{-ikz} \sin\phi, \qquad E_l^i = e^{-ikz} \cos\phi. \tag{5.2.82}$$

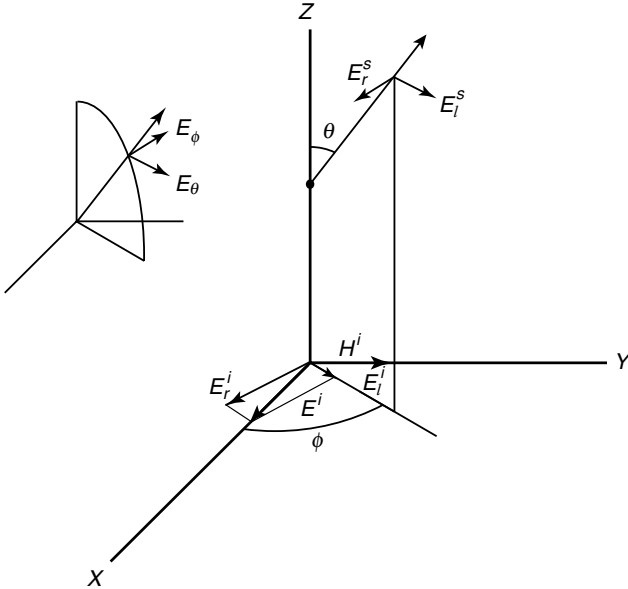
Figure 5.6   Decomposition of the incident (i) and scattered (s) electric vectors into perpendicular (r) and parallel (l) components in Cartesian (x, y, z) and spherical (r, θ, ϕ) coordinates.

Equation (5.2.80) can then be expressed by

$$
\begin{bmatrix} E_l^s \\ E_r^s \end{bmatrix} = \frac{e^{-ikr + ikz}}{ikr} \begin{bmatrix} S_2(\theta) & 0 \\ 0 & S_1(\theta) \end{bmatrix} \begin{bmatrix} E_l^i \\ E_r^i \end{bmatrix},
\tag{5.2.83}
$$

representing the fundamental equation for the study of a light beam scattered by spheres including polarization.

The scattered intensity components in the far-field zone can now be written in terms of the incident intensity components in the form

$$
I_l^s = I_l^i \frac{i_2}{k^2 r^2}, \qquad I_r^s = I_r^i \frac{i_1}{k^2 r^2},
\tag{5.2.84}
$$

where

$$
i_1(\theta) = |S_1(\theta)|^2, \qquad i_2(\theta) = |S_2(\theta)|^2,
\tag{5.2.85}
$$

referred to as the *intensity functions* for the perpendicular and parallel components, respectively. Each of these components of scattered light can be thought of as emanating from that component of the incident light beam polarized in the same direction. The computational problem involved in Lorenz–Mie scattering is to evaluate i_1 and i_2 as functions of the scattering angle, index of refraction, and particle size parameter.

In the far-field zone, we wish to evaluate the reduction of the incident energy due to the absorption and scattering of light by a sphere. For this purpose we may consider incident light polarized linearly in the perpendicular direction. Based on Eq. (5.2.83), the scattered electric field is given by

$$
E_r^s = \frac{e^{-ikr + ikz}}{ikr} S_1(\theta) E_r^i.
\tag{5.2.86}
$$

Further, we consider a point (x, y, z) in the forward direction, i.e., $\theta \approx 0$. In the far field, since $x(y) \ll z$, we have

$$
r = (x^2 + y^2 + z^2)^{1/2} \approx z + \frac{x^2 + y^2}{2z}.
\tag{5.2.87}
$$

Superimposing incident and scattered electric fields in the forward direction yields

$$
E_r^i + E_r^s \approx E_r^i \left[1 + \frac{S_1(0)}{ikz} e^{-ik(x^2 + y^2)/2z} \right].
\tag{5.2.88}
$$

The far-field combined flux density in the forward direction is then proportional to

$$
|E_r^i + E_r^s|^2 \approx |E_r^i|^2 \left\{ 1 + \frac{2}{kz} \, \mathrm{Re} \left[\frac{S_1(0)}{i} e^{-ik(x^2 + y^2)/2z} \right] \right\},
\tag{5.2.89}
$$

where Re[ ] represents the real part of the argument.

We may now integrate the combined flux density over the cross-section area of a sphere whose radius is a to obtain the total power of the combined image as follows:

$$\frac{1}{\left|E_r^i\right|^2} \int\int \left|E_r^i + E_r^s\right|^2 dx\, dy = \pi a^2 + \sigma_e, \tag{5.2.90}$$

where the first term on the right-hand side of Eq. (5.2.90) represents the cross-section area of the sphere. The physical interpretation of the second term, σ_e, is that the total light received in the forward direction is reduced by the presence of the sphere, and that the amount of this reduction is as if an area equivalent to σ_e of the object had been covered up. The double integral over $dx\, dy$ from which σ_e is defined contains two Fresnel integrals. If the limits are assumed to extend to infinity, we obtain

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-ik(x^2+y^2)/2z} dx\, dy = \frac{2\pi z}{ik}. \tag{5.2.91}$$

It follows that the extinction cross section is given by

$$\sigma_e = (4\pi/k^2)\mathrm{Re}[S(0)]. \tag{5.2.92}$$

In the forward direction, we have

$$S_1(0) = S_2(0) = S(0) = \frac{1}{2}\sum_{n=1}^{\infty}(2n+1)(a_n+b_n). \tag{5.2.93}$$

There is only one $S(0)$ because of the symmetry of the forward scattering in which the extinction is independent of the state of polarization of the incident light. It should be noted that Eq. (5.2.92) is valid only when the sphere is isotropic and homogeneous. Furthermore, the extinction efficiency of a sphere with a radius of $r = a$ is defined by

$$Q_e = \frac{\sigma_e}{\pi a^2} = \frac{2}{x^2}\sum_{n=1}^{\infty}(2n+1)\mathrm{Re}(a_n+b_n), \tag{5.2.94}$$

where $x = ka$ is the size parameter defined in Section 1.1.4.

To derive the scattering cross section, we use Eq. (5.2.80). The flux density of the scattered light in an arbitrary direction is given by

$$F(\theta, \phi) = \frac{F_0}{k^2 r^2}\left[i_2(\theta)\cos^2\phi + i_1(\theta)\sin^2\phi\right] \tag{5.2.95}$$

with F_0 representing the incident flux density. The total flux (or power) of the scattered light is, therefore,

$$f = \int_0^{2\pi}\int_0^{\pi} F(\theta, \phi) r^2 \sin\theta\, d\theta\, d\phi, \tag{5.2.96}$$

where $\sin\theta\, d\theta\, d\phi$ is the differential solid angle $d\Omega$, and $r^2\, d\Omega$ denotes the differential area. The scattering cross section is then given by

$$\sigma_s = \frac{f}{F_0} = \frac{\pi}{k^2}\int_0^{\pi}[i_1(\theta) + i_2(\theta)]\sin\theta\, d\theta. \tag{5.2.97}$$

As in the extinction case, we may define the scattering efficiency of a sphere as

$$Q_s = \frac{\sigma_s}{\pi a^2} = \frac{1}{x^2} \int_0^\pi [i_1(\theta) + i_2(\theta)] \sin \theta \, d\theta. \tag{5.2.98}$$

We note the following orthogonal and recurrence properties of the associated Legendre polynomials (see Appendix E):

$$\int_0^\pi \left(\frac{dP_n^1}{d\theta} \frac{dP_m^1}{d\theta} + \frac{1}{\sin^2 \theta} P_n^1 P_m^1 \right) \sin \theta \, d\theta = \begin{cases} 0, & \text{if } n \neq m \\ \dfrac{2n(n+1)}{2n+1} \dfrac{(n+1)!}{(n-1)!}, & \text{if } n = m, \end{cases} \tag{5.2.99}$$

and

$$\int_0^\pi \left(\frac{P_n^1}{\sin \theta} \frac{dP_m^1}{d\theta} + \frac{P_m^1}{\sin \theta} \frac{dP_n^1}{d\theta} \right) \sin \theta \, d\theta = \left[P_n^1(\theta) P_m^1(\theta) \right]_0^\pi = 0. \tag{5.2.100}$$

The scattering efficiency can be evaluated with the help of these two equations to yield

$$Q_s = \frac{2}{x^2} \sum_{n=1}^\infty (2n+1)(|a_n|^2 + |b_n|^2). \tag{5.2.101}$$

Finally, based on the energy conservation principle, the absorption cross section and efficiency of a sphere can be evaluated from the following relationships:

$$\sigma_a = \sigma_e - \sigma_s,$$
$$Q_a = Q_e - Q_s. \tag{5.2.102}$$

For an absorbing sphere, it is convenient to define the index of refraction as $m = m_r - im_i$, with m_r and m_i representing the real and imaginary parts of the refractive index, respectively.

Figure 5.7 shows the scattering efficiency factor Q_s as a function of the size parameter x for a real index of refraction of 1.5 and several values of the imaginary part. For $m_i = 0$, i.e., a perfect reflector, there is no absorption. In this case, $Q_s (= Q_e)$ shows a series of major maxima and minima and ripples. The major maxima and minima are due to interference of light diffracted and transmitted by the sphere, whereas the ripple arises from edge rays that are grazing and traveling around the sphere, spewing off energy in all directions. Q_s (or Q_e) increases rapidly when the size parameter reaches about 5 and approaches an asymptotic value of 2. This implies that a large particle removes from the incident beam exactly twice the amount of light that it can intercept. The light removed from the incident light beam includes the diffracted component, which passes by the particle, plus the light scattered by reflection and refraction within the particle, which will be discussed in the following section. Both the ripples and the major maxima and minima damp out as absorption within the particle increases.
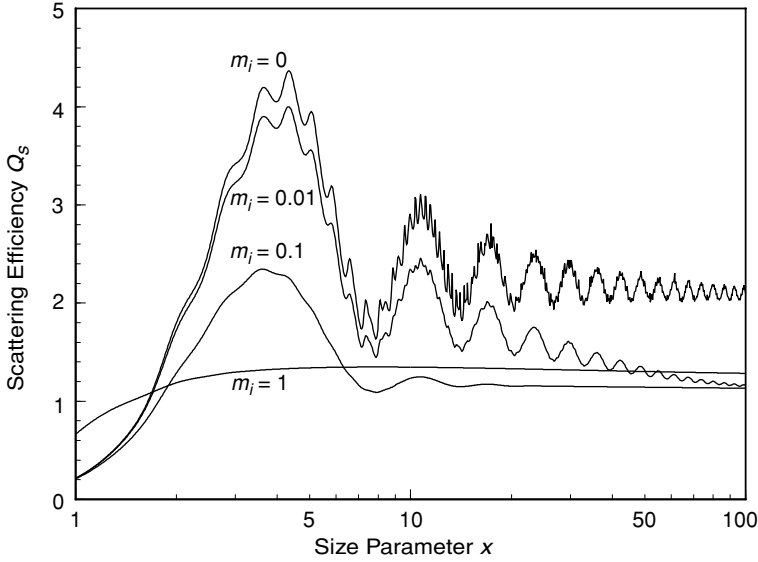
Figure 5.7   Efficiency factor for scattering, Q_s, as a function of the size parameter $x = 2\pi a/\lambda$. The real part of the refractive index used is $m_r = 1.5$, with results shown for four values of the imaginary part m_i.

5.2.4   Scattering Phase Matrix for Spherical Particles

To describe the complete polarization properties of electromagnetic waves, a set of four quantities referred to as the *Stokes parameters*, first introduced by Stokes (1852), has been defined. Because the intensity is proportional to the absolute square of the electric field, we may define, upon neglecting a constant of proportionality, the following four parameters:

$$I = E_l E_l^* + E_r E_r^*, \tag{5.2.103a}$$

$$Q = E_l E_l^* - E_r E_r^*, \tag{5.2.103b}$$

$$U = E_l E_r^* + E_r E_l^*, \tag{5.2.103c}$$

$$V = -i(E_l E_r^* - E_r E_l^*), \tag{5.2.103d}$$

where an asterisk denotes the complex conjugate value and $i = \sqrt{-1}$. I, Q, U, and V are real quantities that satisfy $I^2 = Q^2 + U^2 + V^2$. Further discussion of the Stokes parameters is presented in Section 6.6.1, which deals with multiple scattering of polarized light.

We may express the incident and scattered electric vectors given by Eq. (5.2.83) in terms of their intensity components. Letting the subscript 0 denote the incident

component, we can show that

$$
\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \frac{\mathbf{F}}{k^2 r^2} \begin{bmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{bmatrix}, \tag{5.2.104}
$$

where the matrix

$$
\mathbf{F} = \begin{bmatrix} \frac{1}{2}(M_2 + M_1) & \frac{1}{2}(M_2 - M_1) & 0 & 0 \\ \frac{1}{2}(M_2 - M_1) & \frac{1}{2}(M_2 + M_1) & 0 & 0 \\ 0 & 0 & S_{21} & -D_{21} \\ 0 & 0 & D_{21} & S_{21} \end{bmatrix}, \tag{5.2.105}
$$

and its components are defined by

$$
M_{1,2} = S_{1,2}(\theta) S_{1,2}^*(\theta),
$$
$$
S_{21} = [S_1(\theta) S_2^*(\theta) + S_2(\theta) S_1^*(\theta)]/2,
$$
$$
-D_{21} = [S_1(\theta) S_2^*(\theta) - S_2(\theta) S_1^*(\theta)] i /2. \tag{5.2.106}
$$

The matrix $\mathbf{F}$ is called the *transformation matrix* of the light scattering of a single sphere.

In conjunction with the transformation matrix, we may define a parameter referred to as the *scattering phase matrix* such that

$$
\frac{\mathbf{F}(\theta)}{k^2 r^2} = C\mathbf{P}(\theta). \tag{5.2.107}
$$

The coefficient C can be determined from normalization of the first matrix element in the form

$$
\int_0^{2\pi} \int_0^{\pi} \frac{P_{11}(\theta)}{4\pi} \sin\theta \, d\theta \, d\phi = 1. \tag{5.2.108}
$$

It follows that from Eqs. (5.2.107) and (5.2.108), we obtain

$$
C = \frac{1}{2k^2 r^2} \int_0^{\pi} \frac{1}{2}[M_1(\theta) + M_2(\theta)] \sin\theta \, d\theta = \frac{1}{4k^2 r^2} \int_0^{\pi} [i_1(\theta) + i_2(\theta)] \sin\theta \, d\theta. \tag{5.2.109}
$$

Moreover, from the scattering cross section defined in Eq. (5.2.97), we obtain the coefficient in the form

$$
C = \sigma_s/(4\pi r^2). \tag{5.2.110}
$$

Thus,

$$
\frac{P_{11}}{4\pi} = \frac{1}{2k^2 \sigma_s}(i_1 + i_2) = \frac{1}{2}\left(\frac{P_1}{4\pi} + \frac{P_2}{4\pi}\right), \tag{5.2.111a}
$$

$$
\frac{P_{12}}{4\pi} = \frac{1}{2k^2 \sigma_s}(i_2 - i_1) = \frac{1}{2}\left(\frac{P_2}{4\pi} - \frac{P_1}{4\pi}\right), \tag{5.2.111b}
$$

$$\frac{P_{33}}{4\pi} = \frac{1}{2k^2\sigma_s}(i_3 + i_4), \tag{5.2.111c}$$

$$-\frac{P_{34}}{4\pi} = \frac{i}{2k^2\sigma_s}(i_4 - i_3), \tag{5.2.111d}$$

where

$$i_j = S_j S_j^* = |S_j|^2, \qquad j = 1, 2, \tag{5.2.112a}$$

$$i_3 = S_2 S_1^*, \qquad i_4 = S_1 S_2^*. \tag{5.2.112b}$$

The scattering phase matrix for a single homogeneous sphere is then

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & 0 & 0 \\ P_{12} & P_{11} & 0 & 0 \\ 0 & 0 & P_{33} & -P_{34} \\ 0 & 0 & P_{34} & P_{33} \end{bmatrix}. \tag{5.2.113}$$

In general, if no assumption is made about the shape and position of the scatterer, the scattering phase matrix consists of 16 nonzero elements (see Section 5.4.3). For a single sphere, the independent elements reduce to only four. Graphs of P_{11} (the phase function), P_{12}, P_{33}, and P_{34} as functions of the scattering angle for a real part of the refractive index of 1.5 and a size parameter of 60, are shown in Fig. 5.8. The phase functions of a Lorenz–Mie particle are characterized by strong forward scattering, while large backscattering is noticeable as well. These scattering patterns consist of rapid fluctuation due to interference effects, depending on the size parameter. It is clear that the scattering behavior of a Lorenz–Mie particle differs significantly from that of a Rayleigh molecule (Section 3.3.1). Since a spherical particle is symmetrical with respect to the incident light, its scattering pattern is symmetrical in the intervals $(0°, 180°)$ and $(180°, 360°)$. Thus, we may present the Lorenz–Mie scattering phase function in a polar diagram similar to those displayed in Fig. 1.4. The element P_{33} has the same behavior as the phase function P_{11}, but P_{12} and P_{34} show negative values resulting from the differences in the two intensity functions.

The preceding discussion has been concerned with the scattering of electromagnetic waves by a single homogeneous sphere. We shall now extend the discussion to a sample of cloud or aerosol particles so that practical equations for the calculation of extinction parameters and phase functions can be derived. In what follows, particles are assumed to be sufficiently far from each other and that the distance between them is much greater than the incident wavelength. In this case, the scattering by one particle can be treated independently without reference to that of other particles. Consequently, intensities scattered by a sample of particles may be added without regard to the phase of the scattered waves. This particular scattering phenomenon is referred to as *independent scattering* (see also Section 1.1.4). It is in the context of the independent scattering concept that we present the following discussion.
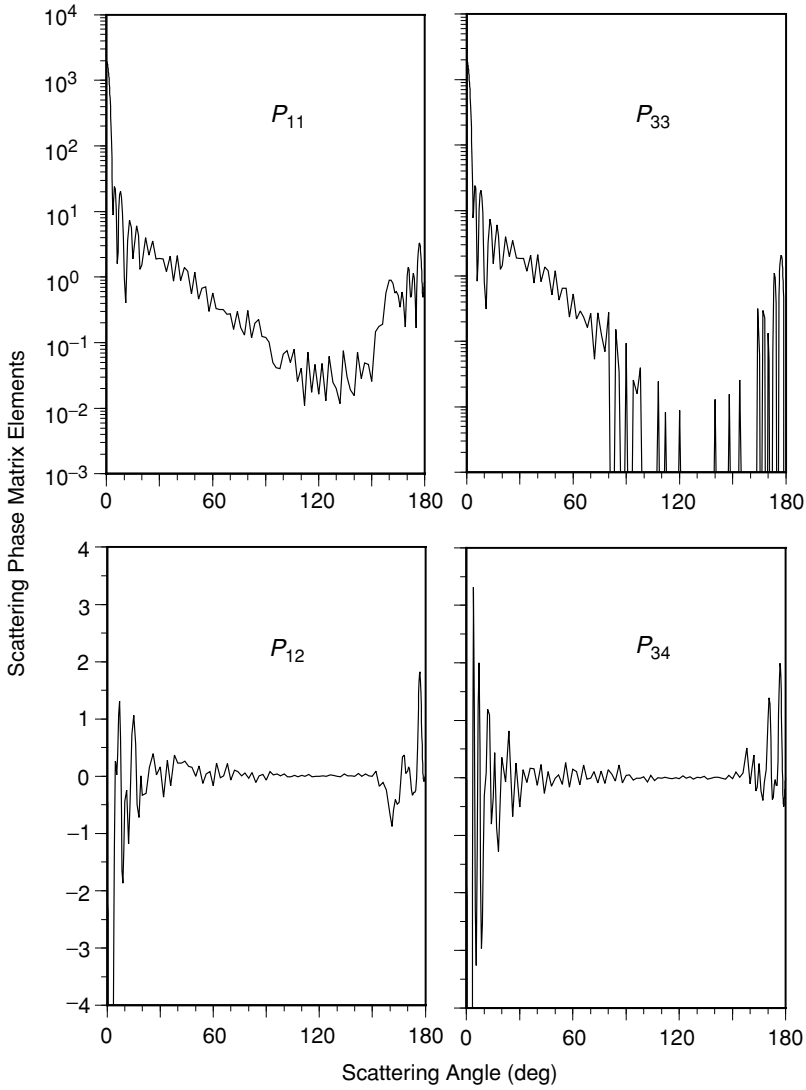
Figure 5.8   Scattering phase matrix elements P_{11}, P_{12}, P_{33}, and P_{34} as functions of the scattering angle for a refractive index $m_r = 1.5$ and a size parameter $x = 60$, computed from the Lorenz–Mie theory.

Consider a sample of cloud particles whose size spectrum can be described by $n(a)$ (in units of cm^{-3} μm^{-1}). Assuming that the size range of particles is from a_1 to a_2, the total number of particles is then given by

$$N = \int_{a_1}^{a_2} n(a)\,da. \tag{5.2.114}$$

In connection with the particle size distribution, the extinction and scattering coefficients (in units of per length) are defined, respectively, as follows:

$$\beta_e = \int_{a_1}^{a_2} \sigma_e(a)n(a)\,da, \tag{5.2.115}$$

$$\beta_s = \int_{a_1}^{a_2} \sigma_s(a)n(a)\,da. \tag{5.2.116}$$

Finally, the single-scattering albedo for a sample of particles is defined by

$$\tilde{\omega} = \beta_s/\beta_e, \tag{5.2.117}$$

which represents the percentage of a light beam that undergoes the scattering event. The remaining part of this section defines the scattering phase matrix for a sample of spherical particles.

Since the scattering phase matrix is a nondimensional physical parameter denoting the scattered intensity and polarization state for a sample of particles in the radius range (a_1, a_2), it is independent of the particle size distribution $n(a)$. Hence, we may rearrange Eq. (5.2.111a) and perform particle size integration to obtain

$$\frac{P_{11}}{4\pi} \int_{a_1}^{a_2} \sigma_s n(a)da = \frac{1}{2k^2} \int_{a_1}^{a_2} [i_1(a) + i_2(a)]n(a)\,da. \tag{5.2.118}$$

From Eq. (5.2.115), we find

$$\frac{P_{11}}{4\pi} = \frac{1}{2k^2\beta_s} \int_{a_1}^{a_2} [i_1(a) + i_2(a)]n(a)\,da. \tag{5.2.119}$$

Similarly, we have

$$\frac{P_{12}}{4\pi} = \frac{1}{2k^2\beta_s} \int_{a_1}^{a_2} [i_2(a) - i_1(a)]n(a)\,da, \tag{5.2.120}$$

$$\frac{P_{33}}{4\pi} = \frac{1}{2k^2\beta_s} \int_{a_1}^{a_2} [i_3(a) + i_4(a)]n(a)\,da, \tag{5.2.121}$$

$$-\frac{P_{34}}{4\pi} = \frac{i}{2k^2\beta_s} \int_{a_1}^{a_2} [i_4(a) - i_3(a)]n(a)\,da. \tag{5.2.122}$$

Note here that the intensity functions $i_j (j = 1, 2, 3, 4)$ are functions of the particle radius a, the index of refraction m, the incident wavelength λ, and the scattering angle θ.

5.3 Geometric Optics

The laws of geometric optics may be used to compute the angular distribution of scattered light when a plane electromagnetic wave is incident on a particle much larger than the wavelength of the incident light. This approach is based on the assumption

that the light may be thought of as consisting of separate localized rays that travel along straight-line paths; it is an asymptotic approach that becomes increasingly accurate in the limit as the size-to-wavelength ratio approaches infinity. Processes involving geometric optics include rays externally reflected by the particle and rays refracted into the particle; the latter rays may be absorbed in the particle or they may emerge from it after possibly suffering several internal reflections. Hence, the total energy scattered and absorbed by the particle is equal to that impinging on the cross section of the particle in reference to the incident beam.

Particles much larger than the incident wavelength also scatter light by means of diffraction, a process by which energy is removed from the light wave passing by the particle. The diffraction is concentrated in a narrow lobe in the forward direction, and like geometrical reflection and refraction, it contains an amount of energy equal to that incident on the cross section of the particle. In the far field, the diffracted component of the scattered light may be approximated by the Fraunhofer diffraction theory. The diffraction pattern depends only upon the shape of the cross section of the particle. The geometric configuration of different contributions to the light scattered by a large sphere was illustrated in Fig. 3.14. In the following, we provide the theoretical foundations of the treatment of geometric optics and diffraction.

5.3.1   Diffraction

The theoretical development of diffraction begins with Babinet's principle, which states that the diffraction pattern in the far field, i.e., Fraunhofer diffraction, from a circular aperture is the same as that from an opaque disk or sphere of the same radius. Let the z axis be in the direction of propagation of the incident light, and let the wave disturbance be sought at a distance P from the geometric aperture A. In reference to Fig. 5.9, the distances from point P to point $O'(x, y)$ on the aperture area and to the origin O are denoted as r and r_0, respectively. Thus, the phase difference of the disturbance at P for waves passing through points O and O' is given by (see Fig. 5.9)

$$\delta = k(r - r_0) = k(x \cos \phi + y \sin \phi) \sin \theta, \tag{5.3.1}$$

where $k = 2\pi/\lambda$, and λ is the wavelength.

In the far field, the light-wave disturbance at P can be derived from the Fraunhofer diffraction theory (Born and Wolf, 1975) and is given by

$$u_p = -\frac{i u_0}{r\lambda} \int\!\!\int_A e^{-ikr} dx\, dy, \tag{5.3.2}$$

where u_0 represents the disturbance in the original wave at point 0 on the plane wave front whose wavelength is λ. Using Eq. (5.3.1), we obtain

$$u_p = -\frac{i u_0}{r\lambda} e^{-ikr_0} \int\!\!\int_A \exp[-ik(x \cos \phi + y \sin \phi) \sin \theta]\, dx\, dy. \tag{5.3.3}$$
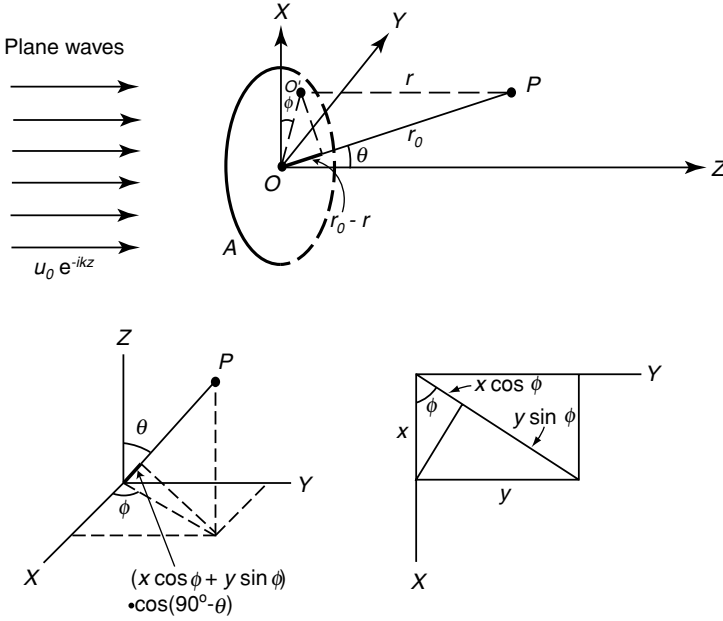
Figure 5.9 Diffraction of plane waves by a circular aperture with a geometrical area A evaluated at position P. The geometrical relationships between the phase difference and the coordinate systems are also shown. All the notations are defined in the text.

For a circular aperture, we may transfer rectangular coordinates to polar coordinates (ρ, ψ) such that $x = \rho \cos \psi$ and $y = \rho \sin \psi$. Thus,

$$u_p = -\frac{i u_0}{r \lambda} e^{-ikr_0} \int_0^a \int_0^{2\pi} \exp[-ik\rho \cos(\psi - \phi) \sin \theta] \, \rho \, d\rho \, d\psi. \qquad (5.3.4)$$

First, we note that the zero-order Bessel function is defined by

$$J_0(y) = \frac{1}{2\pi} \int_0^{2\pi} e^{iy \cos \alpha} d\alpha. \qquad (5.3.5)$$

It follows that

$$u_p = -\frac{i u_0}{r \lambda} e^{-ikr_0} 2\pi \int_0^a J_0(k\rho \sin \theta) \, \rho \, d\rho. \qquad (5.3.6)$$

Second, we also note a well-known recurrence relationship involving Bessel functions given by

$$\frac{d}{dy}[y J_1(y)] = y J_0(y), \qquad (5.3.7)$$

from which we have

$$\int_0^y y' J_0(y')dy' = y J_1(y).$$ (5.3.8)

Thus, from Eqs. (5.3.6) and (5.3.8), we obtain

$$u_p = -\frac{iu_0}{r\lambda}e^{-ikr_0}A\frac{2J_1(x\sin\theta)}{x\sin\theta},$$ (5.3.9)

where the geometric shadow area $A = \pi a^2$ and the size parameter $x = ka$. Hence, the scattered intensity in terms of the incident intensity $I_0 = |u_0|^2$ is given by

$$I_p = |u_p|^2 = I_0\frac{i_p}{k^2r^2},$$ (5.3.10)

where the angular intensity function for diffraction analogous to the Lorenz–Mie theory for a single sphere is defined by

$$i_p = \frac{x^4}{4}\left[\frac{2J_1(x\sin\theta)}{x\sin\theta}\right]^2.$$ (5.3.11)

It is clear that diffraction depends only on the particle size parameter and is independent of the index of refraction.

Figure 5.10 shows a plot of $D^2 = [2J_1(y)/y]^2$ versus y. It has a principal maximum of 1 at $y = 0$ (i.e., $\theta = 0$) and oscillates with gradually diminishing amplitude as y
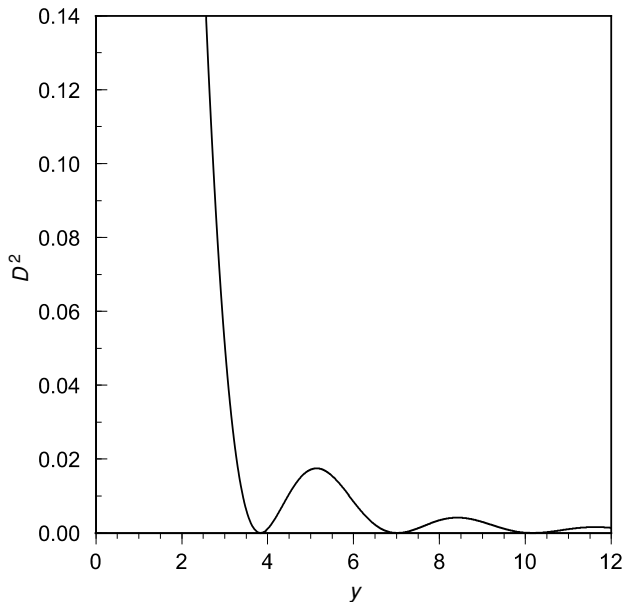

Figure 5.10   The relative diffraction pattern (see text for the definition) as a function of the parameter $y = x\sin\theta$.

increases. When $J_1(y) = 0$, $D^2 = 0$, giving the minima of the diffraction pattern. The positions of the maxima are defined by values of y that satisfy

$$\frac{d}{dy}[J_1(y)/y] = 0. \tag{5.3.12}$$

Table 5.2 lists these maxima and minima. The minima, or dark rings, can be approximated by

$$y = ka \, \sin \theta = (n + 0.22)\pi, \quad n = 1, 2, \ldots, \tag{5.3.13}$$

or

$$\sin \theta = (n + 0.22)\lambda/(2a). \tag{5.3.14}$$

The first maximum at $y = 0$ usually is obscured by the finite size of the source. Thus, the first observable maximum diffraction ring is for $y = 5.136$.

The diffraction theory for a single sphere developed above can be used to explain the optical phenomenon known as a *corona*. A corona is associated with the illumination frequently observed near the sun, the moon, or other luminous objects when they are seen through a mist or thin cloud. They are usually observed in the form of circles, or near circles, concentric with the luminous body. A corona is usually very bright and of a white or bluish-white color with a reddish or brownish tinge. The colors are diluted with a great deal of white light. As many as four corona rings have been recorded, though only the first ring has been frequently observed around the sun and the moon. A corona may be produced by thin clouds when such clouds are composed of particles of almost equal size, a condition called *monodisperse*. Applying the preceding diffraction theory, one may evaluate the angular positions of the corona if the wavelength of the visible sunlight and the mean particle size are known. Based on Eq. (5.3.14), it is evident that red color, having the longer wavelength in the visible, must be seen in the outer ring of the corona with blue and green colors inside the ring. Also note that the angular width of the corona depends on the diameter of the particles involved.

Table 5.2

The First Few Maxima and Minima of the
Diffraction Pattern

y	D^2	Max or min
0	1	Max
3.832	0	Min
5.136	0.0175	Max
7.016	0	Min
8.417	0.0042	Max
10.174	0	Min
11.620	0.0016	Max

5.3.2   Geometric Reflection and Refraction

When a plane wave impinges on a boundary separating two homogeneous media of different optical properties, it is split into two waves: a transmitted wave entering into the second medium and a reflected wave propagating back into the first medium. From the part of the wave that hits the surface plane, we may isolate a narrow beam that is much smaller than the surface. Such a beam is commonly called a *ray* as in geometric optics. Let v_1 and v_2 be the velocities of propagation in the two media ($v_1 > v_2$), and let θ_i and θ_t be the angles corresponding to the incident and refracted waves. Referring to Fig. 5.11, we find (see also the subsection on geometric optics in Section 3.3.2.)

$$\sin \theta_i / \sin \theta_t = v_1/v_2 = m, \tag{5.3.15}$$

where m is the index of refraction for the second medium with respect to the first medium. This is Snell's law, which relates the incident and refracted angles through the index of refraction.

Let $\mathbf{E}^i$ be the electric vector of the incident field. As shown in Fig. 5.11, the components of the incident electric field vector perpendicular (r) and parallel (l) to the

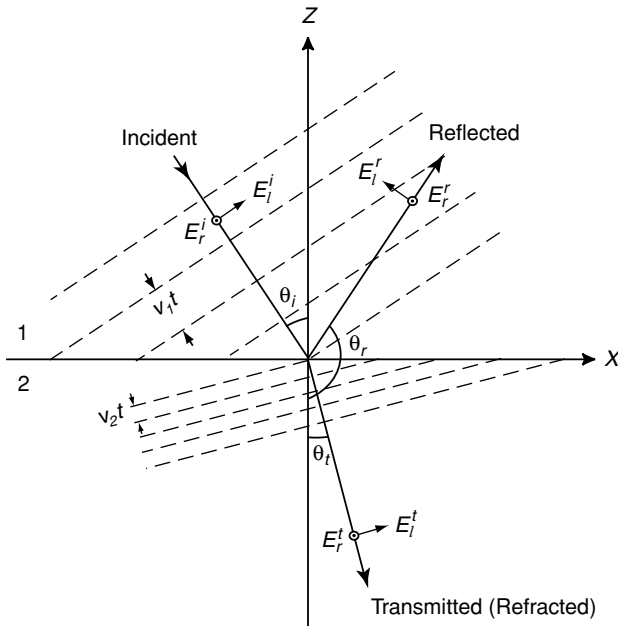**Figure 5.11**   Illustration of the reflection and refraction of a plane wave. The choice of the positive directions for the parallel components (l) of the electric vectors is indicated in the diagram. The perpendicular components are at right angles into the plane of reference. $v_1 t$ and $v_2 t$ denote the distance in media 1 and 2, respectively, and other notations are defined in the text. See also Fig. 3.14a.

plane containing the incident and refracted fields mapped in rectangular coordinates are given by

$$E_x^i = -E_l^i \cos \theta_i,$$
$$E_y^i = E_r^i,$$
$$E_z^i = E_l^i \sin \theta_i. \tag{5.3.16}$$

From the Maxwell equations, the relationship between the electric and magnetic vectors can be shown to be $\mathbf{H} = \sqrt{\varepsilon}\, \mathbf{a} \times \mathbf{E}$, or $\mathbf{E} = -\sqrt{1/\varepsilon}\, \mathbf{a} \times \mathbf{H}$, where $\mathbf{a}$ is a unit vector in the direction of propagation. Thus, the components of the magnetic vector are as follows ($\mu \approx 1$, $m = \sqrt{\varepsilon}$):

$$H_x^i = -E_r^i \cos \theta_i m_1,$$
$$H_y^i = -E_l^i m_1,$$
$$H_z^i = E_r^i \sin \theta_i m_1, \tag{5.3.17}$$

where m_1 is the refractive index of the first medium with respect to vacuum.

Likewise, if $\mathbf{E}^t$ and $\mathbf{E}^r$ denote the transmitted (refracted) and reflected electric vectors, respectively, we find the following relationships:

$$E_x^t = -E_l^t \cos \theta_t, \qquad H_x^t = -E_r^t \cos \theta_t m_2,$$
$$E_y^t = E_r^t, \qquad H_y^t = -E_l^t m_2,$$
$$E_z^t = E_l^t \sin \theta_t, \qquad H_z^t = E_r^t \sin \theta_t m_2, \tag{5.3.18}$$

$$E_x^r = -E_l^r \cos \theta_r, \qquad H_x^r = -E_r^r \cos \theta_r m_1,$$
$$E_y^r = E_r^r, \qquad H_y^r = -E_l^r m_1,$$
$$E_z^r = E_l^r \sin \theta_r, \qquad H_z^r = E_r^r \sin \theta_r m_1, \tag{5.3.19}$$

where m_2 is the refractive index of the second medium with respect to vacuum, and $\theta_r = 180° - \theta_i$.

The boundary conditions require that the tangential components of $\mathbf{E}$ and $\mathbf{H}$ be continuous. Hence, we must have

$$E_x^i + E_x^r = E_x^t, \qquad H_x^i + H_x^r = H_x^t,$$
$$E_y^i + E_y^r = E_y^t, \qquad H_y^i + H_y^r = H_y^t. \tag{5.3.20}$$

On substituting all the electric and magnetic components into Eq. (5.3.20), we obtain the following four relationships:

$$\cos \theta_i \left(E_l^i - E_l^r\right) = \cos \theta_t E_l^t, \tag{5.3.21a}$$

$$E_r^i + E_r^r = E_r^t, \tag{5.3.21b}$$

$$m_1 \cos \theta_i \left(E_r^i - E_r^r \right) = m_2 \cos \theta_t E_r^t, \tag{5.3.21c}$$

$$m_1 \left(E_l^i + E_l^r \right) = m_2 E_l^t. \tag{5.3.21d}$$

On the basis of these equations, solutions of the electric components of the reflected and transmitted waves in terms of the incident wave are given by

$$E_l^r = R_l E_l^i, \quad E_l^t = T_l E_l^i, \quad E_r^r = R_r E_r^i, \quad E_r^t = T_r E_r^i, \tag{5.3.22}$$

where the amplitude coefficients are

$$R_r = \frac{\cos \theta_i - m \cos \theta_t}{\cos \theta_i + m \cos \theta_t}, \quad R_l = \frac{m \cos \theta_i - \cos \theta_t}{m \cos \theta_i + \cos \theta_t}, \tag{5.3.23a}$$

$$T_r = \frac{2 \cos \theta_i}{\cos \theta_i + m \cos \theta_t}, \quad T_l = \frac{2 \cos \theta_i}{m \cos \theta_i + \cos \theta_t}, \tag{5.3.23b}$$

with $m = m_2/m_1$, the refractive index of the second medium with respect to the first medium. Equations (5.3.23a) and (5.3.23b) are referred to as *Fresnel formulas* (Fresnel, 1823). When absorption is involved, the amplitude coefficients become much more complicated, but they can be derived by means of a straightforward mathematical analysis (see Exercise 5.8). Interested readers may wish to refer to Section 5.4.1 for further discussion.

From the Poynting vector, $\mathbf{S} = (c/4\pi) \, \mathbf{E} \times \mathbf{H}$, and the relation between $\mathbf{E}$ and $\mathbf{H}$ denoted previously, we find that the flux density $|\mathbf{S}| = (c/4\pi)\sqrt{\varepsilon}|\mathbf{E}|^2 (\mu = 1)$. Thus, the amount of energy incident, reflected, and transmitted on a unit area of the boundary per unit time is

$$F^i = |\mathbf{S}^i| \cos \theta_i = (c/4\pi)m_1 |\mathbf{E}^i|^2 \cos \theta_i, \tag{5.3.24a}$$

$$F^r = |\mathbf{S}^r| \cos \theta_r = (c/4\pi)m_1 |\mathbf{E}^r|^2 \cos \theta_r, \tag{5.3.24b}$$

$$F^t = |\mathbf{S}^t| \cos \theta_t = (c/4\pi)m_2 |\mathbf{E}^t|^2 \cos \theta_t. \tag{5.3.24c}$$

Therefore, the reflected and transmitted portions of the energy in two polarization components, with respect to the incident energy, are proportional to $R_{r,l}^2$ and $T_{r,l}^2 m \cos \theta_t / \cos \theta_i$, respectively. It can be proven that $R_{r,l}^2 + T_{r,l}^2 m \cos \theta_t / \cos \theta_i = 1$, in fulfillment of the energy conservation principle. It follows that the transmitted (or refracted) parts of the energy can be written simply as $(1 - R_{r,l}^2)$.

Consider now a large sphere and let $p = 0$ for the external reflection, $p = 1$ for two refractions, and $p \geq 2$ for internal reflections. We may define the amplitude coefficients as follows:

$$\varepsilon_1 = R_r \qquad\qquad\qquad\text{for} \quad p = 0,$$

$$\varepsilon_1 = (1 - R_r^2)^{1/2}(-R_r)^{p-1}(1 - R_r^2)^{1/2} \quad \text{for} \quad p \geq 1, \tag{5.3.25}$$

where $-R_r$ denotes the amplitude coefficient for an internal reflection. These definitions also apply to the index $l(2)$ of the other polarization component.

Next, we discuss the effect of curvature on the reflected and refracted intensity. We consider a finite pencil of light characterized by $d\theta_i$ and $d\phi$, with ϕ being the azimuthal

angle. Let I_0 denote the incident intensity of the light pencil plane-polarized in one of the two main directions. The flux of energy contained in this pencil is, therefore, $I_0 a^2 \cos\theta_i \sin\theta_i d\theta_i d\phi$, where a is the radius of the sphere. This flux of energy is divided by successive reflection and refraction. The emergent pencil spreads into a solid angle, $\sin\theta d\theta d\phi$, at a large distance, r, from the sphere. As a result, the scattered intensity is given by

$$I_r = \frac{\varepsilon_1^2 I_0 a^2 \cos\theta_i \sin\theta_i d\theta_i d\phi}{r^2 \sin\theta d\theta d\phi}. \qquad (5.3.26)$$

The pencil of light emergent from the sphere is characterized by a small range $d\theta$ around the scattering angle θ. In reference to Fig. 5.12, the total deviation from the original direction is

$$\theta' = 2(\theta_i - \theta_t) + 2(p-1)(\pi/2 - \theta_t). \qquad (5.3.27a)$$

The scattering angle defined in the interval $(0, \pi)$ may be expressed by

$$\theta' = 2\pi n - q\theta, \qquad (5.3.27b)$$

where n is an integer and $q = +1$ or -1. Hence,

$$\frac{d\theta}{d\theta_i} = \left|\frac{d\theta'}{d\theta_i}\right| = 2 - 2p\frac{\cos\theta_i}{m\cos\theta_t}. \qquad (5.3.28)$$

Moreover, we define the divergence factor due to the curvature effect in the form

$$D = \frac{\cos\theta_i \sin\theta_i}{\sin\theta d\theta/d\theta_i}. \qquad (5.3.29)$$

Thus, from Eq. (5.3.26), which is also valid for index l, we obtain

$$I_{r,l} = \frac{I_0}{k^2 r^2} x^2 \varepsilon_{1,2}^2 D. \qquad (5.3.30)$$

In comparison with the Lorenz-Mie scattering theory, we find

$$i_{1,2} = x^2 \varepsilon_{1,2}^2 D. \qquad (5.3.31)$$

The rainbow is probably the best-known phenomenon of atmospheric optics. It has inspired art and mythology in all peoples and has been a challenge for mathematical physicists. We see a rainbow in the sky usually on summer afternoons after a rain shower is over. Rainbows are produced by the geometric reflections of the sun's rays within raindrops. The sun's rays undergo minimum deviation within the drops and generate the maximum intensity at a specific angle that is much stronger than that at neighboring angles.

To evaluate the angles at which rainbows are formed, we return to the equation denoting the total deviation from the original direction. The minimum deviation of a bundle of rays may be found by differentiating Eq. (5.3.27a) with respect to the incident angle and setting the result equal to zero. Thus,

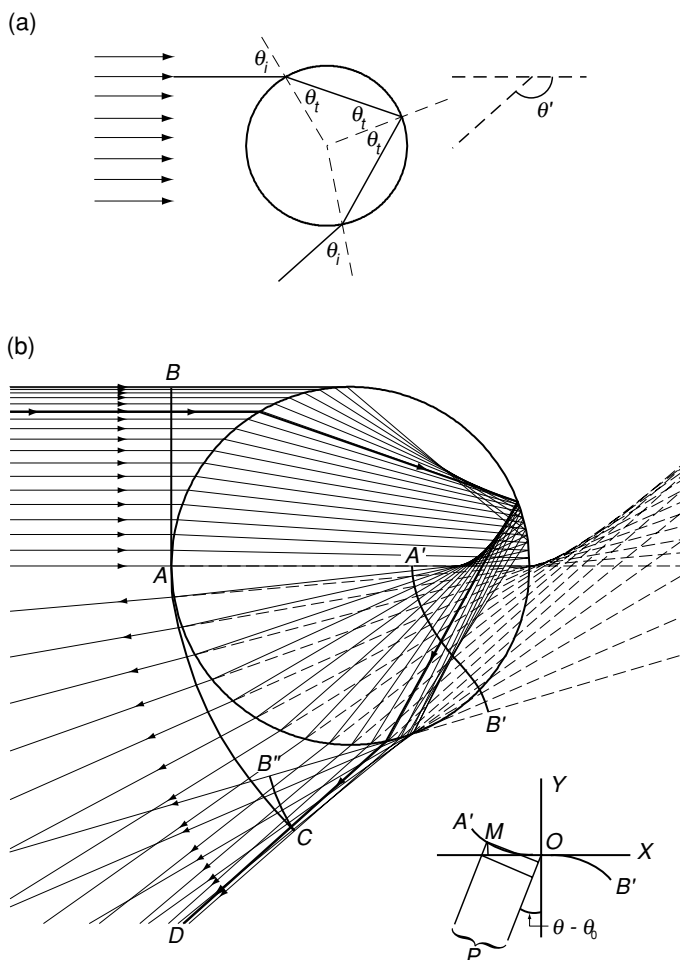$$\frac{d\theta'}{d\theta_i} = 0 = 2\left(1 - p\frac{d\theta_t}{d\theta_i}\right). \qquad (5.3.32)$$

(a)

(b)

Figure 5.12   (a) Geometrical reflection and refraction by a sphere and the definition of the deviation angle θ'. The incident and refracted angles are denoted by θ_i and θ_t, respectively. (b) Refraction and reflection of light rays with the plane wave front denoted by AB. The heavy line represents the ray that undergoes minimum deviation, referred to as the Descartes ray. The line $A'B'$ denotes the outgoing wave front near the Descartes ray. The line ACB'' denotes the outgoing wave front outside the sphere. The attached diagram shows the phase difference at the wave front $A'B'$, as described in the text (data provided by Yoshihide Takano of the University of California, Los Angeles.)

Furthermore, differentiating Snell's law leads to

$$\frac{d\theta_t}{d\theta_i} = \frac{\cos\theta_i}{m\cos\theta_t}. \tag{5.3.33a}$$

It follows that

$$p\cos\theta_i = m\cos\theta_t. \tag{5.3.33b}$$

On eliminating the refracted angle θ_t from Eq. (5.3.33b) and Snell's equation, we obtain the incident angle at which the minimum deviation takes place as follows:

$$\cos^2 \theta_i = (m^2 - 1)/(p^2 - 1), \qquad p \geq 2. \qquad (5.3.34)$$

Once the refractive index has been given, we may evaluate the incident angle corresponding to the minimum deviation for a given p. The refracted angle and the scattering angle may also be subsequently calculated from Snell's law and Eqs. (5.3.27a) and (5.3.27b), respectively.

Table 5.3 lists the incident and scattering angles for rainbows with various indices of refraction. Other than rainbows, minimum deviation can also occur at $\theta = 180°$, known as *glory*. The condition for producing a glory associated with the $p = 2$ rays is that the refractive index must be between $\sqrt{2}$ and 2. Since the refractive index for water drops is about 1.33, glory cannot be produced by means of geometric ray-tracing. The fact that glory is often observed in clouds reveals one of the greatest discrepancies between geometric optics and the Lorenz–Mie theory. It has been suggested that the glory phenomenon is produced by the backscattering from edge rays apparently connected with surface waves generated on a sphere. Primary ($p = 2$) and secondary ($p = 3$) rainbows frequently are observed in the atmosphere. Because of variation of the index of refraction of water with respect to the incident visible wavelengths, various color sequences are produced. Exercise 5.7 requires the calculation of the color position for primary and secondary rainbows. When rainbows are very pronounced, supernumerary rainbows often become visible. They are produced by interference phenomena accompanying refraction of the light in the raindrop and cannot be explained by the geometric ray-tracing approach.

The preceding analysis only defines the angles at which rainbows can be produced. However, geometric optics cannot define rainbow intensity. It is evident from Eq. (5.3.30) that when minimum deviations occur, the intensity approaches infinity. The geometric optics approximation assumes that the wave fronts near any point are sufficiently characterized by their normals and local radii of curvature. Such an approximation breaks down near the rainbow. The next highest approximation can be

Table 5.3

Incident (θ_i) and Scattering (θ) Angles for the Glory and Rainbows

	Glory for $p = 2$		Rainbow for $p = 2$		Rainbow for $p = 3$		Rainbow for $p = 4$		Rainbow for $p = 5$	
m	θ_i	θ	θ_i	θ	θ_i	θ	θ_i	θ	θ_i	θ
1.10	—	—	75°	84°	81°	139°	83°	170°	85°	119°
1.33	—	—	60	137	72	130	77	43	80	42
1.45	87°	180°	53	152	68	102	74	4	78	92
1.50	83	180	50	157	67	93	73	9	77	109
1.54	79	180	47	161	66	86	72	19	76	121
1.75	60	180	34	173	59	58	68	60	73	175
2.00	33	180	0	180	52	35	63	94	69	140

obtained by using a cubic wave front, which leads to the Airy theory. It should be noted, however, that the intensity results derived from geometric optics are accurate in the vicinity of the rainbow angles, and with appropriate extrapolations, intensity patterns at the rainbow angles may be approximated. We shall demonstrate this later when comparing the ray optics and Lorenz–Mie approaches.

In the following, we introduce the Airy theory for rainbows. Since a raindrop is assumed to be spherical, it suffices to use only a single plane containing the center of the drop and the luminous object, and to trace rays incident on one quadrant of the intersection circle. In reference to Fig. 5.12b, let AB be the wave front of parallel incident rays above the ray that passes through the center of the drop, referred to as the axial ray, and consider rays that undergo only one internal reflection. The heavy line denotes the course of the ray of minimum deviation, referred to as the Descartes ray, for a water drop having an index of refraction of 4/3. Because deviations of the incident rays that are between the Descartes and axial rays are greater than that of the Descartes ray, their exits must lie between those two specific rays. Likewise, the rays that are located between the Descartes and edge rays must also have more deviations than the Descartes ray. As a consequence, they emerge from the drop beyond the Descartes ray and eventually come between this ray and the axial ray. The internally reflected light rays are therefore diffuse and weakened, except near the direction of minimum deviation, and are confined to the region between this direction and the axial ray. The wave front is now described by ACB''.

The outgoing wave front near the Descartes ray is in the form of $A'B'$. Through tedious geometric analysis and numerous approximations (e.g., see Humphreys, 1954), it can be shown that the wave front in Cartesian coordinates is related by the cubic equation in the form

$$y = hx^3/(3a^2), \tag{5.3.35a}$$

where a is the radius of the drop and

$$h = \frac{(p^2 + 2p)^2}{(p + 1)^2(m^2 - 1)}\left[\frac{(p + 1)^2 - m^2}{m^2 - 1}\right]^{1/2}. \tag{5.3.35b}$$

Equation (5.3.35b) represents a curve closely coincident with the portion of the wave front by which rainbows are produced.

To evaluate the intensity and its variation with angular distance from the ray of minimum deviation, we consider the diagram depicted in Fig. 5.12b, and let O be the point of inflection of the outgoing cubic wave front near a drop. Let θ_0 and θ be the deviation angles for the Descartes ray and the neighboring rays, respectively, and let P be a distant point in the direction $\theta - \theta_0$ from the Descartes ray. We find that the phase difference of the disturbance at P for waves passing through points O and M is given by

$$\delta = k[x\sin(\theta - \theta_0) - y\cos(\theta - \theta_0)]$$
$$= k\left[x\sin(\theta - \theta_0) - \frac{h}{3a^2}x^3\cos(\theta - \theta_0)\right], \tag{5.3.36}$$

where $k = 2\pi/\lambda$. The amplitude of the wave disturbance u_p is then proportional to the integration of all the possible vibrations due to phase differences along the x axis as

$$u_p \sim \int_{-\infty}^{\infty} \exp\left\{-ik\left[x\sin(\theta - \theta_0) - \frac{h}{3a^2}x^3\cos(\theta - \theta_0)\right]\right\} dx. \qquad (5.3.37a)$$

It suffices to use the cosine representation, and if we let

$$(2h/3a^2\lambda)x^3\cos(\theta - \theta_0) = t^3/2, \qquad (2x/\lambda)\sin(\theta - \theta_0) = zt/2, \qquad (5.3.38a)$$

the amplitude is now given by

$$u_p \sim 2\left[\frac{3a^2\lambda}{4h\cos(\theta - \theta_0)}\right]^{1/3} f(z). \qquad (5.3.37b)$$

The intensity $I = u_p^2$, where the *rainbow integral* due to Airy is defined by

$$f(z) = \int_0^{\infty} \cos\frac{\pi}{2}(zt - t^3)dt. \qquad (5.3.38a)$$

From Eq. (5.3.38a), we obtain

$$z^3 = \frac{48a^2}{h\lambda^2}\frac{\sin^3(\theta - \theta_0)}{\cos(\theta - \theta_0)}, \qquad (5.3.38b)$$

which, for small values of $\theta - \theta_0$, is proportional to $(\theta - \theta_0)^3$. Thus, for a small departure from the Descartes ray, we have

$$z \approx \left(\frac{48a^2}{h\lambda^2}\right)^{1/3}(\theta - \theta_0). \qquad (5.3.38c)$$

Table 5.4 lists maximum and minimum values of z and $f^2(z)$ for a given wavelength and drops of a definite size. Note that the first maximum (main rainbow) does not coincide with $z = 0$, the geometric position of the primary rainbow ($\theta = 138.0°$). Also note that the absolute intensity derived from the Airy theory may be obtained by comparing the result from the geometric optics method with the value of $f^2(z)$ for large z.

The validity of the Airy theory for rainbows was investigated by van de Hulst (1957), who indicated that it is an accurate approach when $x \geq 5000$ and the deviation from the rainbow angle $|\varepsilon| \leq 0.5°$. Nussenzveig (1979, and the reference papers therein) presented a more advanced mathematical analysis based on the asymptotic behavior of the exact Lorenz–Mie solution for the physical explanation of rainbows and glory. The two scattering functions denoted in Eq. (5.2.78) can be expressed in the form

$$S_j(x, \theta) = S_{j,0}(x, \theta) + \sum_{p=1}^{N} S_{j,p}(x, \theta) + \text{remainder}, \quad j = 1, 2, \qquad (5.3.39a)$$

Table 5.4

Maxima and Minima of the Rainbow Integral

Maxima			Minima	
Number	z	$f^2(z)$	Number	z
1	1.0845	1.005	1	2.4955
2	3.4669	0.615	2	4.3631
3	5.1446	0.510	3	5.8922
4	6.5782	0.450	4	7.2436
5	7.8685	0.412	5	8.4788
6	9.0599	0.384	6	9.6300
7	10.1774	0.362	7	10.7161
8	11.2364	0.345	8	11.7496
9	12.2475	0.330	9	12.7395
10	13.2185	0.318	10	13.6924

where the first term on the right-hand side of Eq. (5.3.39a) is associated with the direct reflection from the surface of a sphere, and the second term is associated with transmission after $(p - 1)$ internal reflections at the surface. Equation (5.3.39a) is referred to as the *Debye expansion*. The terms in this equation can be transformed in a complex domain based on the Poisson sum formula, referred to as the *Watson transformation,* in the form

$$\sum_{\ell=0}^{\infty} \varphi(\Lambda, \mathbf{r}) = \sum_{m=-\infty}^{\infty} (-1)^m \int_0^{\infty} \varphi(\Lambda, \mathbf{r}) \exp(2im\pi\Lambda) d\Lambda, \qquad (5.3.39b)$$

where φ is a certain function, $\Lambda = \ell + 1/2$, ℓ is the order of expansion in the scattering functions, and $\mathbf{r}$ is a position vector. In this manner, the real rays are mapped into a complex domain, allowing the expression of the scattering functions $S_j(j = 1, 2)$ in terms of the Fresnel coefficients (geometric optics) and the Airy integral, which is called the *complex angular momentum theory* of the rainbow. This theory can be used to generate the intensity peaks produced in rainbow angles, including polarization, associated with internal reflections.

In the preceding discussion, we pointed out that because of its limitations, the principle of geometric optics cannot be utilized to explain the formation of the glory feature, which is often observed on mountaintops and from aircraft, because of the required backscattering geometry. Following the complex angular momentum theory, the tangential rays may be thought of as composing complex angular momentum that can tunnel into the regions not accessible to geometric ray-tracing as forms of surface waves on the edge of a sphere. These rays can then undergo critical refraction to the inside at a point, total reflection at a suitable point inside, and refraction to the outside at an opposite symmetric point. Subsequently, these rays travel as surface waves in the backscattering direction to the observer, giving rise to the glory phenomenon. It turns out that the glory is produced by the ray path undergoing two shortcuts associated with the $p = 2$ Debye term in Eq. (5.3.39b).

5.3.3   Geometric Optics, Lorenz–Mie Theory, and Representative Results

In order to compare the scattering results derived from the geometric optics approximation with those from the Lorenz–Mie theory, we define the gain G relative to an isotropic scatterer. This gain is defined as the ratio of the scattered intensity to the intensity that would be found in any direction if the particle scattered the incoming energy isotropically. Thus, the average gain over the entire solid angle is 1 such that

$$\frac{1}{4\pi} \int_{4\pi} G_{1,2}(\theta)d\Omega = 1. \tag{5.3.40}$$

Isotropic scattering implies that the incident energy $I_0\pi a^2$ to a sphere with radius a is uniformly distributed over its surface $4\pi r^2$. Consequently,

$$I_{1,2}(\theta) = \frac{I_0\pi a^2}{4\pi r^2} G_{1,2}(\theta). \tag{5.3.41}$$

From Eq. (5.2.84), we have

$$G_{1,2} = 4i_{1,2}/x. \tag{5.3.42a}$$

In a similar manner, the gain due to diffraction can be written as follows:

$$G^f = 4i_p/x, \tag{5.3.42b}$$

where i_p is given in Eq. (5.3.11). The gain due to diffraction is the same for the perpendicular and parallel components.

The total gain caused by diffraction and geometric reflection and refraction can now be expressed by

$$G^t_{1,2} = G^f + \sum_{p=0}^{N} G^p_{1,2}, \tag{5.3.43}$$

where the geometric reflection and refraction are represented by the index p, including external reflection $(p = 0)$, refraction $(p = 1)$, and internal reflection $(p \geq 2)$.

The foregoing discussion neglects the different contributions to the scattered intensity caused by the phase interferences produced by various ray components. In the case of large particles, phase interferences give rise to rapidly oscillating intensities as a function of the scattering angle as shown in Fig 5.8. However, if particles are randomly located and separated by distances much larger than the incident wavelength, a condition referred to as *polydisperse,* the intensities from separate particles may be added without regard to the phase because the numerous maxima and minima are then lost in the integration over particle size. Hence, it is reasonable to neglect the phase altogether in adding the intensities for diffraction, reflection, and refraction for a sample of large particles of various sizes.

Figure 5.13 compares phase function P_{11} and the degree of linear polarization computed from the Lorenz–Mie theory to the corresponding results from the geometric optics approach for typical refractive indices of 1.33 and 1.50, representing water drops and aerosols, respectively, in the visible spectrum. The degree of linear
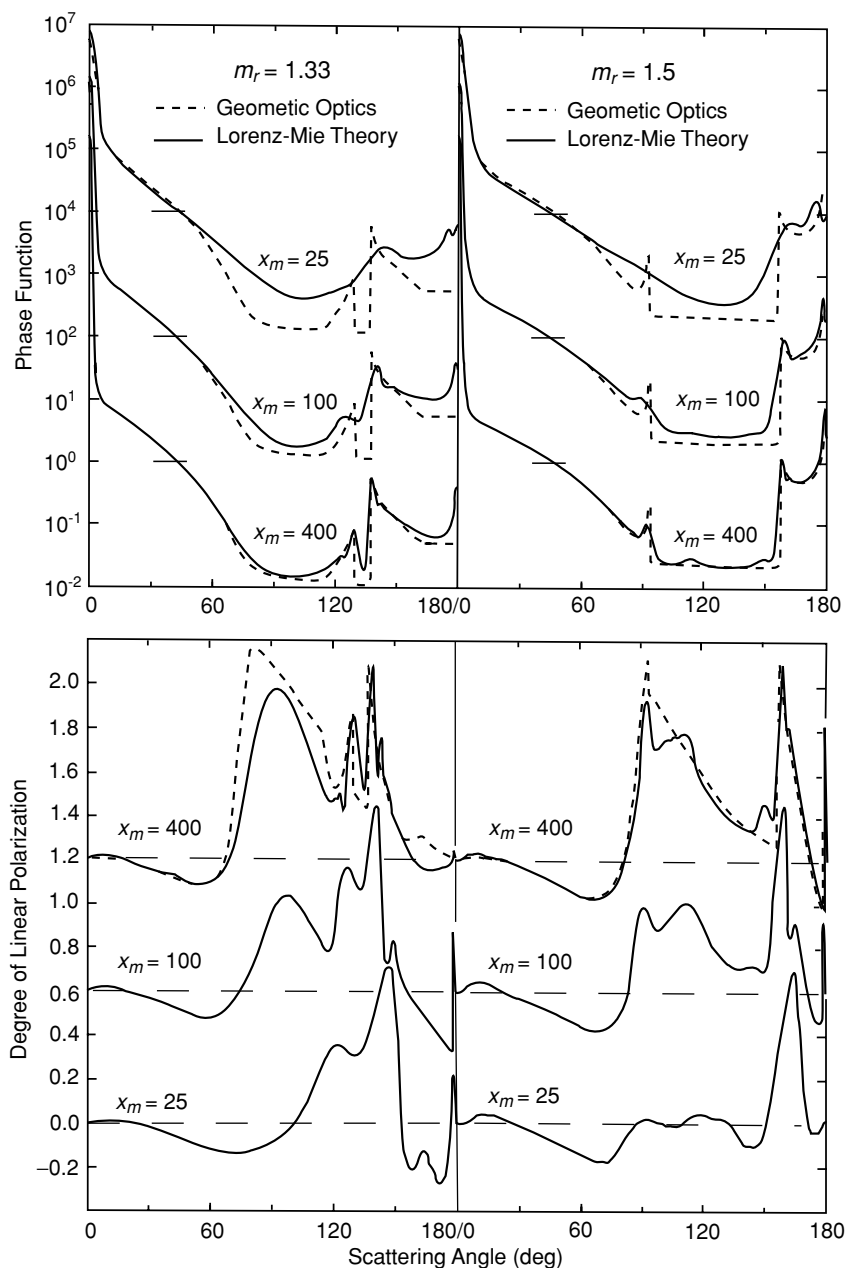
Figure 5.13   Comparison of geometric optics and the Lorenz–Mie theory for phase function P_{11} and degree of linear polarization. Two refractive indices are shown along with three size distributions. The vertical scale applies to the lowermost curves. The other curves are successively displaced upward by factors of 10^2 and shifts of 0.6 for the phase function and the degree of linear polarization, respectively (after Liou and Hansen, 1971).

polarization is defined by $-Q/I$ from Eq. (5.2.104); for incident unpolarized sunlight (see Section 6.6.1 for further discussion), it is given by $-P_{12}/P_{11}$. Lorenz–Mie calculations were made for size parameters of 25, 100, and 400. The droplet size distribution employed in the Lorenz–Mie and diffraction calculations to smooth out rapid oscillating patterns is a gamma function given by

$$n(x) = Cx^6 \exp(-6x/x_m), \qquad (5.3.44)$$

where C is an arbitrary constant and x_m is the modal parameter. There is close agreement between geometric optics and the Lorenz–Mie theory when the size parameter is as large as 400. An exception is the glory feature for $m_r = 1.33$, which, as discussed earlier, does not occur in geometric optics results. Most of the discrepancies and their variations with the size parameter can be qualitatively understood in terms of the increasing inapplicability of the localization principle for decreasing size parameters. This causes the light in the individual features to be blurred over a wider range of angles than predicted by geometric optics. The secondary rainbow is quite smooth at $x_m = 100$ and is lost at $x_m = 25$, while the primary rainbow is still visible. The number of rainbows visible in the intensity pattern thus gives some indication of the particle size. The rainbow, in addition to being smoothed out, tends to move away from its geometric optics location as the size parameter decreases. For Lorenz–Mie scattering with $x_m = 400$, the small secondary peaks on the less steep side of the rainbows are supernumerary bows, which are caused by interference phenomena and hence are not rendered by geometric optics, which neglects phase interference effects. There is also a small but noticeable discrepancy in the diffraction peak. The higher value for Lorenz–Mie scattering could be produced by surface waves that scatter in the forward direction. The lower figure compares the degree of linear polarization, which contains much stronger imprints of most of the features occurring in the scattered light, such as rainbows, supernumerary bows, the glory, and external reflections, all of which produce positive polarization.

In Fig. 5.14, we illustrate comparisons of theoretical and experimental scattering phase function and polarization patterns. These scattering patterns are derived from measurements of a dense water cloud in a cold chamber, utilizing a He–Ne (0.6328 μm) laser light. The curves depicted in this figure comprise five successive nephelometer scans (10°–175°) that were each normalized at 10° scattering angles and then averaged, with the standard deviations of the averages shown as vertical bars. The measured cloud droplet size, using a continuous-impactor–replicator device, displayed a modal diameter of 2 μm and a maximum diameter of 10 μm. This size spectrum was fitted with a zeroth order log-normal distribution in the form

$$n(a) = \exp\left[-(\log a - \log a_m)^2/(2\sigma_o^2)\right]\Big/\sqrt{2\pi}\sigma_0 a, \qquad (5.3.45)$$

where a_m ($= 2\ \mu$m) denotes the modal diameter, and σ_0 ($= 0.275$) is the geometric mean standard deviation. Lorenz–Mie scattering calculations were performed employing this size distribution and the results were compared to the measured data.
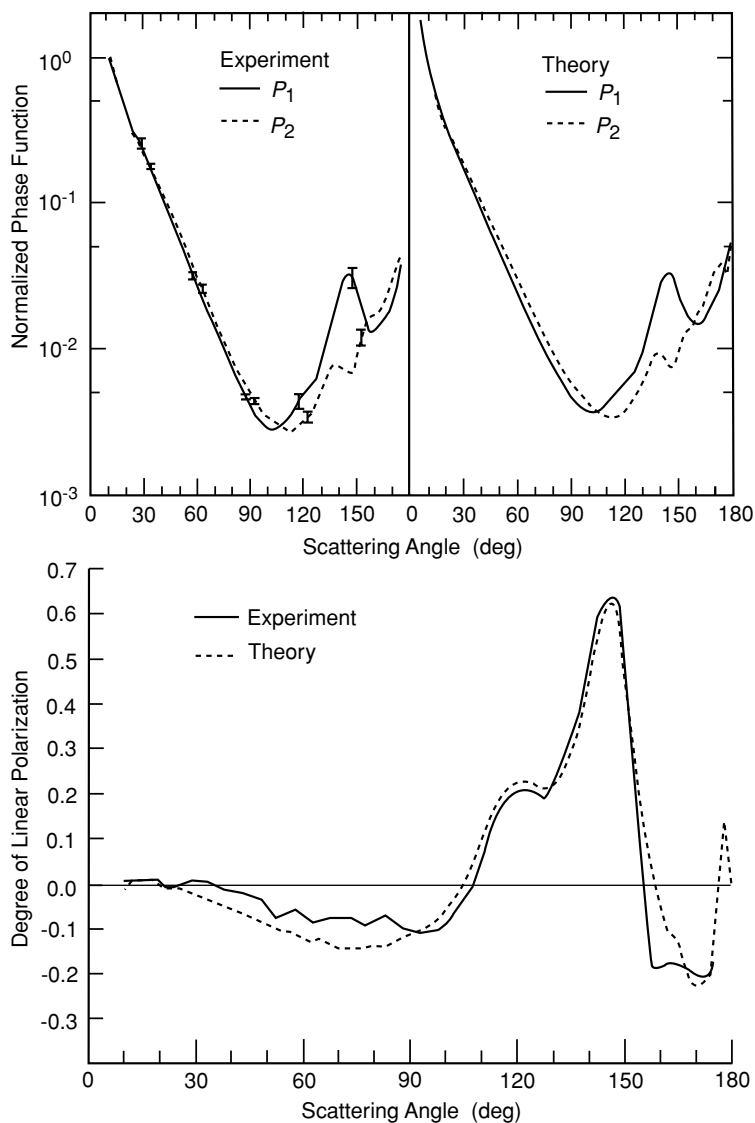
Figure 5.14   Comparison of the normalized phase function components P_1 and P_2 [see Eq. (5.2.111a) for the definitions] and the degree of linear polarization from water cloud experiments and Lorenz–Mie scattering theory, using the same cloud droplet size distribution (after Sassen and Liou, 1979).

The close agreement between the theory and measurements is illustrated in Fig. 5.14. Because of the relatively small size of cloud droplets (the modal size parameter $x_m \approx$ 20), both theory and measurements reveal that the secondary cloudbow is absent, and the primary cloudbow reaches a maximum at about $146°$. Close agreement is

also found for the linear polarization pattern, especially in the vicinity of the primary cloudbow where large positive polarization values are shown.

Lorenz–Mie scattering calculations can now routinely be carried out for any size parameters based on more efficient formulations and vector structures of computers (see, e.g., Wiscombe, 1980). In the following, we present some representative results for phase functions, linear polarizations, average extinction cross sections, and single-scattering albedos for water clouds computed from the Lorenz–Mie theory. Using the droplet size distribution denoted in Eq. (5.3.44), the phase functions and the degree of linear polarization patterns for the 0.63, 1.6, 3.7, and 10 μm wavelengths, typical of the image channels, e.g., the Advanced Very High Resolution Radiometer (AVHRR) onboard NOAA satellites, are presented in Fig. 5.15. At 0.63 μm the feature shown at a scattering angle of about 140° is the well-known primary rainbow that arises from rays undergoing one internal reflection in the cloud droplets. This cloudbow decreases in strength at 1.6 μm because of higher absorption and the smaller size parameters involved. At 0.63 μm, the minor maximum located at about 120° is the secondary rainbow, which is produced by rays undergoing two internal reflections. The maximum that occurs at about 180° for 0.63 μm is the glory pattern. The glory feature moves toward smaller scattering angles at 1.6 μm. At 3.7 and 10 μm, the rainbow and glory features vanish because of the significant absorption by cloud droplets. Also
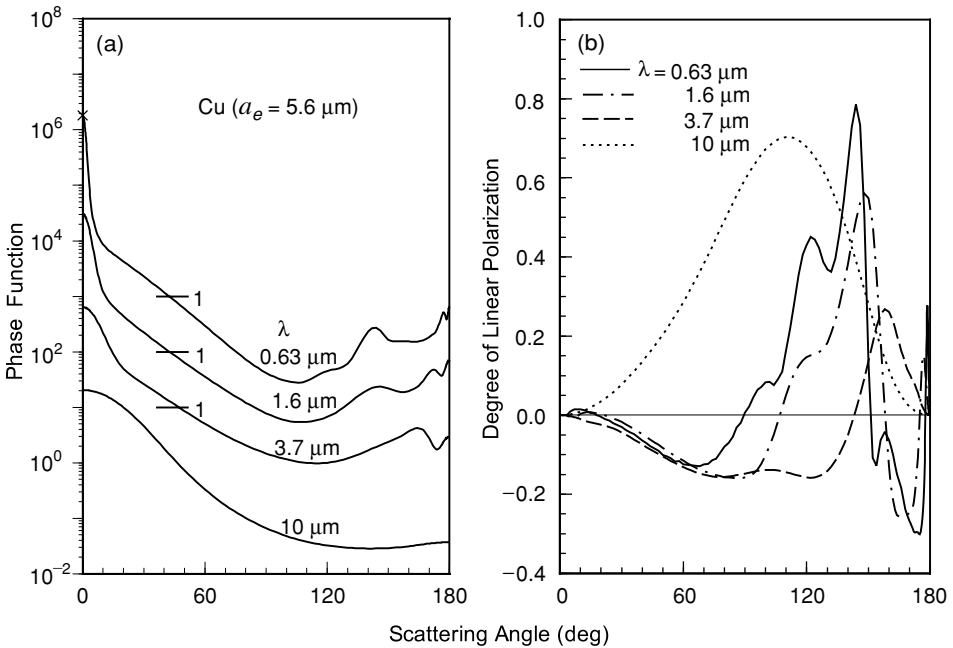
Figure 5.15 Phase function (a) and degree of linear polarization (b) for incident wavelengths of 0.63, 1.6, 3.7, and 10 μm involving the droplet size distribution representative of cumulus and stratus clouds. For the phase function, the vertical scale applies to the lowest curve, while the upper curves are displayed upward by a factor of 10. The symbol × denotes the diffraction peak for the 0.63 μm wavelength.

noted is the reduction of the Fraunhofer diffraction peak when the size parameter is decreased. The polarization patterns for 0.63 μm contain many strong imprints of scattered light in the directions of the rainbow and glory angles. The positions of the primary rainbows for 1.6 and 3.7 μm shift to larger scattering angles because of the effect of the refractive index and size parameter. For the 10 μm wavelength, scattering is primarily associated with external reflection, which produces a maximum positive polarization at about 90° scattering angle.

Figure 5.16 displays the average extinction cross section, single-scattering albedo, and asymmetry factor for typical cumulus clouds with a droplet size distribution
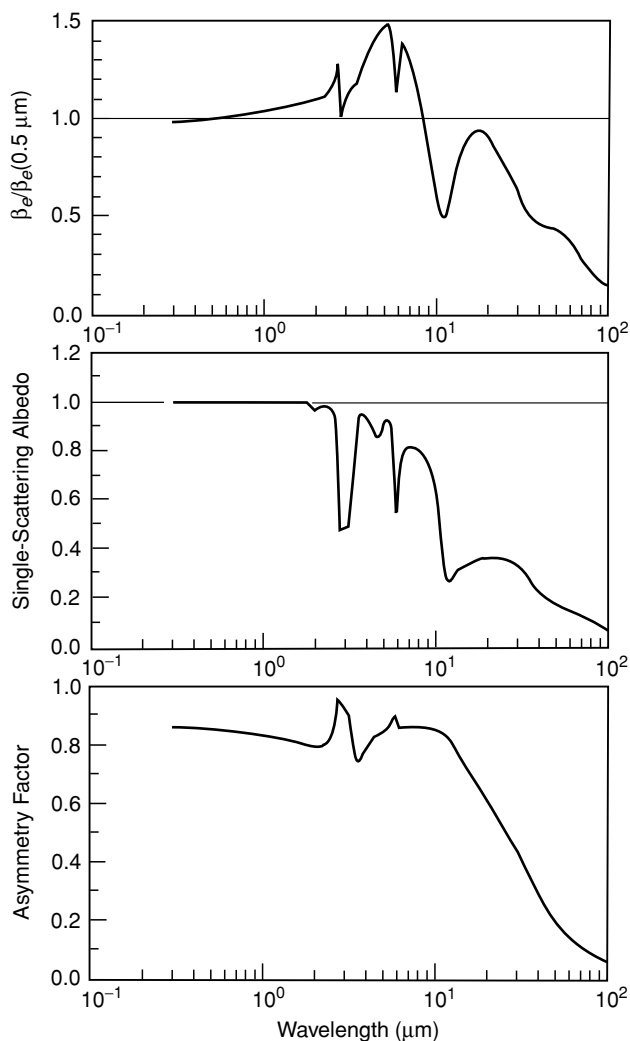

Figure 5.16   Normalized extinction coefficient [$\beta_e(0.5\ \mu$m) $= 41.8$ km^{-1}], single-scattering albedo, and asymmetry factor for the droplet size distribution of typical low cumulus and stratus clouds, covering the wavelengths from 0.3 to 100 μm.

similar to the gamma function denoted in Eq. (5.3.44). The average extinction cross section is normalized with respect to its value at 0.5 μm. In the visible wavelengths, the average size parameter is sufficiently large that the extinction efficiency is approximately equal to 2, based on the optical theorem described previously. Extinction efficiency varies with the size parameter and refractive index. The single-scattering albedo generally resembles the maximum and minimum patterns in the imaginary refractive index for water. Three large minima occur at about 3, 6, and 10 μm. For wavelengths longer than 10 μm, the size parameter effect becomes more important. The asymmetry factor remains about the same, with values of about 0.82–0.86, for wavelengths up to about 10 μm, with exceptions at about 3 and 6 μm; at these wavelengths, diffraction predominates and produces strong forward scattering. Because of the reduction of the size parameter, the asymmetry factor decreases drastically for wavelengths longer than 10 μm.

5.4   Light Scattering by Ice Crystals: A Unified Theory

The scattering of light by spheres can be solved by the exact Lorenz-Mie theory presented in Section 5.2, and computations can be performed for the size parameters that are practical for atmospheric applications. However, an exact solution for the scattering of light by nonspherical ice crystals covering all sizes and shapes that occur in the earth's atmosphere, as displayed in Section 5.1, does not exist in practical terms. It is unlikely that one specific method can be employed to resolve all of the scattering problems associated with nonspherical ice crystals. In the following, we present a unified theory for light scattering by ice crystals by means of a combination of the geometric optics and finite-difference time domain methods.

5.4.1   Geometric Optics for Ice Crystals

As discussed in Section 5.3, the principles of geometric optics are the asymptotic approximations of the fundamental electromagnetic theory, valid for light scattering computations involving a target whose dimension is much larger than the incident wavelength. The geometric optics method has been employed to identify the optical phenomena occurring in the atmosphere, such as halos, arcs, and rainbows. In addition, it is currently the only practical approach for the solution of light scattering by large nonspherical particles. In this section, we present the conventional and improved approaches, the methodology dealing with absorption in the context of geometric ray tracing, and the numerical implementation of the Monte Carlo method.

5.4.1.1   CONVENTIONAL APPROACH

When the size of an ice crystal is much larger than the incident wavelength, we may consider a light beam as consisting of a bundle of separate parallel rays that undergo reflection and refraction outside and inside the ice crystal with propagation directions determined by the Snell law at the surface. The total field is assumed to consist of the diffracted rays and the reflected and refracted rays. The diffracted rays pass around

the ice crystal, while the rays impinging on the ice crystal undergo local reflection and refraction, referred to as *Fresnelian interaction*. The energy that is carried by the diffraction and the Fresnelian rays is assumed to be the same as the energy that is intercepted by the ice crystal cross section projected along the incident direction. The intensity of the scattered light within the small scattering-angle interval $\Delta\theta$ in the scattering direction θ in the far field can be computed from the summation of the intensity contributed by each individual ray emerging from the direction between $\theta + \Delta\theta/2$ and $\theta - \Delta\theta/2$. It is usually assumed that the interference is smoothed out when the particles are randomly oriented. In this case, the extinction efficiency (the ratio of the extinction cross section to the average projected area of the particle) of the scatterer is 2. On the basis of Babinet's principle, diffraction by a scatterer may be regarded as that by an opening on an opaque screen perpendicular to the incident light, which has the same geometric shape as the projected cross section of the scatterer. The Fraunhofer diffraction approach can then be employed to compute the diffraction component for hexagonal ice particles (Cai and Liou, 1982).

In the geometric ray-tracing method, the directions of the rays are determined first. In reference to Fig. 5.17, these rays can be defined by the following unit vectors:

$$\mathbf{e}_p^r = \mathbf{x}_p - 2(\mathbf{x}_p \cdot \mathbf{n}_p)\mathbf{n}_p, \quad p = 1, 2, 3, \dots, \tag{5.4.1a}$$
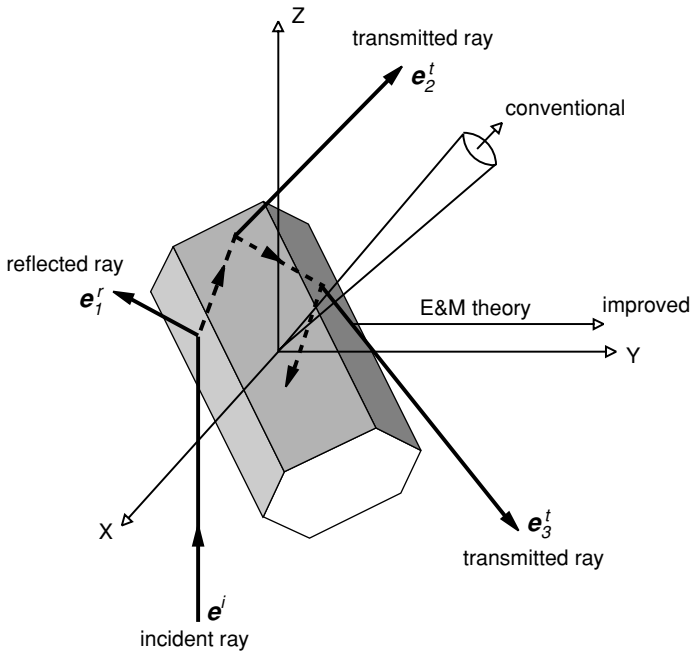

Figure 5.17   Geometry of ray tracing involving an ice crystal in three-dimensional space. Conventional and improved methods are also indicated in the diagram. The vector **e** is a directional cosine associated with incident, reflected, and transmitted rays.

$$\mathbf{e}_p^t = \frac{1}{m_p}\left\{\mathbf{x}_p - (\mathbf{x}_p \cdot \mathbf{n}_p)\mathbf{n}_p - \left[m_p^2 - 1 + (\mathbf{x}_p \cdot \mathbf{n}_p)^2\right]^{1/2}\mathbf{n}_p\right\}, \quad p = 1, 2, 3, \ldots,$$

$$(5.4.1b)$$

$$\mathbf{x}_p = \begin{cases} \mathbf{e}^i & p = 1, \\ \mathbf{e}_1^t, & p = 2, \\ \mathbf{e}_{p-1}^r, & p \geq 3, \end{cases} \qquad (5.4.1c)$$

where the superindices *i, r,* and *t* denote incident, reflected, and transmitted rays; $m_p = m$ for $p = 1$ and $m_p = 1/m$ for $p > 1$, with m being the refractive index; and $\mathbf{n}_p$ denotes a unit vector normal to the surface. When $m_p^2 < 1 - (\mathbf{x}_p \cdot \mathbf{n}_p)^2$, total reflection occurs so that the refracted ray does not exist. The electric fields for two polarization components associated with the rays can be computed from the Fresnel formulas [see Eq. (5.3.23)]. By summing the energy of the rays that emerge in a given direction within a preset small scattering-angle interval, the phase function can be obtained for this part of the scattering process. Let the normalized phase functions for the parts of reflection and refraction, and diffraction be P_{11}^r and P_{11}^d, respectively. The normalized phase function is then $P_{11} = (1 - f_d)P_{11}^r + f_d P_{11}^d$, where $f_d = 1/[2\tilde{\omega}(1 - f_\delta)]$ is the delta transmission associated with the $0°$ refraction produced by two parallel prismatic faces, and $\tilde{\omega}$ is the single-scattering albedo, which can be determined from the absorption of individual rays and the constant extinction efficiency.

5.4.1.2   IMPROVED GEOMETRIC OPTICS APPROACH

The laws of geometric optics are applicable to light scattering by a particle whose size is much larger than the incident wavelength so that rays can be localized. In addition to the requirement of the localization principle, the conventional geometric ray-tracing technique assumes that the energy attenuated by the scatterer may be decomposed into equal extinction from diffraction and Fresnel rays. Moreover, the Fraunhofer diffraction formulation used in geometric ray-tracing does not account for the vector property of the electromagnetic field. Finally, direct calculations of the far field by ray-tracing will produce a discontinuous distribution of the scattered energy, such as the delta transmission noted by Takano and Liou (1989a).

To circumvent a number of shortcomings in the conventional geometric optics approach, an improved method has been developed (Yang and Liou, 1995, 1996a). The concept is simple in that the energies determined from geometric ray-tracing at the particle surface are collected and mapped to the far field based on the exact internal geometric ray-tracing. This differs from the conventional approach, which collects energies produced by geometric reflections and refractions directly at the far field through a prescribed solid angle.

The tangential components of the electric and magnetic fields on a surface S that encloses the scatterer can be used to determine the equivalent electric and magnetic currents for the computation of the scattered far field on the basis of the *electromagnetic equivalence theorem* (Schelkunoff, 1943). In this theorem, the electromagnetic field detected by an observer outside the surface would be the same as if the scatterer

were removed and replaced by the equivalent electric and magnetic currents given by

$$\mathbf{J} = \mathbf{n}_s \times \mathbf{H}, \tag{5.4.2a}$$

$$\mathbf{M} = \mathbf{E} \times \mathbf{n}_s, \tag{5.4.2b}$$

where $\mathbf{n}_s$ is the outward unit vector normal to the surface. For the far-field region, we have

$$\mathbf{E}^s(\mathbf{r}) = \frac{e^{ikr}}{ikr} \frac{k^2}{4\pi} \left(\frac{\mathbf{r}}{r}\right) \int \int_S \left[\mathbf{M}(\mathbf{r}') + \left(\frac{\mathbf{r}}{r}\right) \times \mathbf{J}(\mathbf{r}')\right] \exp\left(-ik\mathbf{r} \cdot \frac{\mathbf{r}'}{r}\right) d^2\mathbf{r}', \tag{5.4.3}$$

where $\mathbf{r}/r$ denotes the scattering direction, $\mathbf{r}$ is the reference position vector, $\mathbf{r}'$ is the position vector of the source point, k is the wavenumber, and $i = \sqrt{-1}$. The far-field solution can also be determined by a volume integral involving the internal field.

By means of geometric ray-tracing, the electric field on the surface of a particle can be evaluated after the successive application of Fresnel reflection and refraction coefficients parallel and perpendicular to a defined reference plane at the point of interaction taking into account the path length in the three-dimensional geometry. If an ice crystal shape is of great complexity, as with an aggregate, the surface can be defined as a cubic box so that the computation of the electric field can be conducted on a regularly shaped surface. The electric field of the illuminated and shadowed sides can be defined as follows:

$$\mathbf{E}(\mathbf{r}) = \begin{cases} \mathbf{E}_a(\mathbf{r}) + \mathbf{E}_b(\mathbf{r}), & \mathbf{r} \in \text{ illuminated side,} \\ \mathbf{E}_b(\mathbf{r}), & \mathbf{r} \in \text{ shadowed side,} \end{cases} \tag{5.4.4a}$$

where

$$\mathbf{E}_a(\mathbf{r}) = \mathbf{E}_i(\mathbf{r}) + \mathbf{E}_1^r(\mathbf{r}), \tag{5.4.4b}$$

$$\mathbf{E}_b(\mathbf{r}) = \sum_{p=2}^{\infty} \mathbf{E}_p^t(\mathbf{r}). \tag{5.4.4c}$$

In these equations, $\mathbf{E}_i$ is the incident electric field, $\mathbf{E}_1^r$ is the electric field for external reflection, and $\mathbf{E}_p^t$ are the electric fields produced by two refractions and internal reflections ($p \geq 2$). Because the transverse electromagnetic wave condition is implied in ray-tracing, the magnetic field for each reflection and refraction can be obtained from

$$\mathbf{H}_p^{r,t}(\mathbf{r}) = \mathbf{e}_p^{r,t} \times \mathbf{E}_p^{r,t}(\mathbf{r}), \qquad \text{for } \mathbf{r} \in \text{ outside the particle.} \tag{5.4.5}$$

In practice, the mapping of the near field solution to the far field can be done in its entirety for $\mathbf{E}_a$ in Eq. (5.4.4b). However, for $\mathbf{E}_b$ in Eq. (5.4.4c), the mapping is done ray by ray and the results will include the diffraction pattern. Full account of phase interferences is taken in this mapping process in the determination of the phase function.

In accord with the conservation principle for electromagnetic energy concerning the Poynting vector (Jackson, 1975), the extinction and absorption cross sections of the particle can be derived as follows:

$$\sigma_e = \text{Im} \left\{ \frac{k}{|\mathbf{E}_i|^2} \int \int \int_V (\varepsilon - 1) \mathbf{E}(\mathbf{r}') \cdot \mathbf{E}_i^*(\mathbf{r}') d^3 \mathbf{r}' \right\}, \tag{5.4.6a}$$

$$\sigma_a = \frac{k}{|\mathbf{E}_i|^2} \int \int \int_V \varepsilon_i \mathbf{E}(\mathbf{r}') \cdot \mathbf{E}^*(\mathbf{r}') d^3 \mathbf{r}', \tag{5.4.6b}$$

where the asterisk denotes the complex conjugate, ε_i is the imaginary part of the permittivity, and V is the particle volume.

Finally, when the ray-tracing technique is applied to obtain the surface field, one must properly account for the area elements from which the externally reflected and transmitted localized waves make a contribution to the surface field. If the cross section of the incident localized wave is $\Delta\sigma_i$, the area of the particle surface for external reflection is

$$\Delta\sigma_1^r = -\Delta\sigma_i (\mathbf{n}_i \cdot \mathbf{e}^i)^{-1}. \tag{5.4.7a}$$

For the transmitted rays, the area is given by

$$\Delta\sigma_p^t = -\Delta\sigma_i (\mathbf{n}_1 \cdot \mathbf{e}_1^t) \left[(\mathbf{n}_1 \cdot \mathbf{e}^i)(\mathbf{n}_p \cdot \mathbf{e}_p^t) \right]^{-1}, \qquad p = 2, 3, 4, \dots, \tag{5.4.7b}$$

where all unit vectors have been defined in Eqs. (5.4.la)–(5.4.lc). The radius of the cross section of a ray should be on the order of k^{-1} so that the phase change over the ray cross section is not significant and permits taking proper account of the phase interference of the localized waves by using the phase information at the centers of the rays. Because the phase variation over the ray cross section can be neglected, the numerical results are not sensitive to the shape of the ray cross sections. We may use a circular shape in the calculations.

5.4.1.3   ABSORPTION EFFECTS IN GEOMETRIC OPTICS

The geometric optics approach that has been used in the past generally assumes that the effect of absorption within the particle on the propagating direction of a ray can be neglected so that the refracted angle and the ray path length can be computed from Snell's law and the geometry of the particle. This approach is correct if absorption is weak, as in the case of ice and water at most solar wavelengths. For strong absorption cases, rays refracted inside the particle are almost totally absorbed, so that the geometric optics method can also be used to compute diffraction and external reflection as long as the particle size is much larger than the incident wavelength. Although the preceding argument is physically correct in the limits of weak and strong absorption, we shall consider the general absorption effect in the context of geometric optics based on the fundamental electromagnetic wave theory. Note that the effect of the complex refractive index on geometric optics has been formulated only for the Fresnel coefficients (Stratton, 1941; Born and Wolf, 1975).
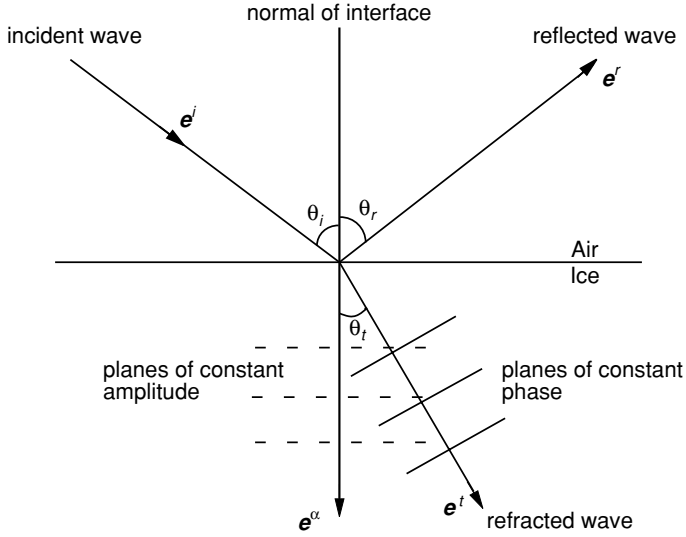
Figure 5.18 Geometric ray-tracing in a medium with absorption. The planes of constant amplitude of the refracted wave are parallel to the interface, whereas the direction of the phase propagation for the inhomogeneous wave inside the medium is determined via Snell's law.

Consider the propagation of the incident wave from air into ice (Fig. 5.18). The wave vectors associated with the incident and reflected waves are real because these waves, which are outside the ice medium, must have the same properties. However, the wave vector of the refracted wave is complex; this is referred to as the inhomogeneity effect. These wave vectors can be presented by

$$\mathbf{k}_i = k\,\mathbf{e}^i, \quad \mathbf{k}_r = k\,\mathbf{e}^r, \quad \mathbf{k}_t = k_t\,\mathbf{e}^t + ik_\alpha\,\mathbf{e}^\alpha, \tag{5.4.8}$$

where $\mathbf{e}^i$, $\mathbf{e}^r$, $\mathbf{e}^t$, and $\mathbf{e}^\alpha$ are unit vectors; the subscripts i, r, and t denote the incident, reflected, and refracted waves respectively; $k = 2\pi/\lambda$, in which λ is the wavelength in air; and k_t and k_α are two real parameters that determine the complex wave vector of the refracted wave. For nonabsorptive cases, k_α is zero. The corresponding electric vectors can be expressed by

$$\mathbf{E}_i(\mathbf{r}, t) = \mathbf{A}_i \exp[i(k\mathbf{r} \cdot \mathbf{e}^i - \omega t)], \tag{5.4.9a}$$

$$\mathbf{E}_r(\mathbf{r}, t) = \mathbf{A}_r \exp[i(k\mathbf{r} \cdot \mathbf{e}^r - \omega t)], \tag{5.4.9b}$$

$$\mathbf{E}_t(\mathbf{r}, t) = \mathbf{A}_t \exp[i(k_t\mathbf{r} \cdot \mathbf{e}^t + ik_\alpha\mathbf{r} \cdot \mathbf{e}^\alpha - \omega t)], \tag{5.4.9c}$$

where $\mathbf{A}_i$, $\mathbf{A}_r$, and $\mathbf{A}_t$ are the amplitudes and ω is the circular frequency. Further, we define the following parameters:

$$N_r = \frac{k_t}{k}, \quad \tilde{N}_i = \frac{k_\alpha}{k}. \tag{5.4.10}$$

At the interface of the two media, at which the position vector is denoted as $\mathbf{r}_s$, the phases of the wave vibration must be the same for the incident, reflected, and refracted waves. Thus, from Eqs. (5.4.8) and (5.4.10), we obtain

$$\mathbf{e}^i \cdot \mathbf{r}_s = \mathbf{e}^r \cdot \mathbf{r}_s = N_r(\mathbf{e}^t \cdot \mathbf{r}_s) + i\tilde{N}_i(\mathbf{e}^\alpha \cdot \mathbf{r}_s). \tag{5.4.11}$$

Because the wave vectors for the incident and reflected waves are real, we must have

$$\mathbf{e}^i \cdot \mathbf{r}_s = \mathbf{e}^r \cdot \mathbf{r}_s = N_r(\mathbf{e}^t \cdot \mathbf{r}_s), \qquad \mathbf{e}^\alpha \cdot \mathbf{r}_s = 0. \tag{5.4.12}$$

Based on the geometry defined by Eq. (5.4.12), a generalized form of the Snell law can be derived and is given by

$$\sin\theta_i = \sin\theta_r, \quad \sin\theta_t = \frac{\sin\theta_i}{N_r}, \tag{5.4.13}$$

where θ_i, θ_r, and θ_t denote the incident, reflected, and refracted angles, respectively (Fig. 5.18). The vector $\mathbf{e}_\alpha$ in Eq. (5.4.12) is normal to the interface of the two media. It follows that the planes of constant amplitude of the refracted wave are parallel to the interface. To determine N_r and N_i, we use the electric field of the refracted wave, which must satisfy the wave equation in the form

$$\nabla^2 \mathbf{E}_t(\mathbf{r}, t) - \frac{(m_r + im_i)^2}{c^2} \frac{\partial^2 \mathbf{E}_t(\mathbf{r}, t)}{\partial t^2} = 0, \tag{5.4.14}$$

where c is the speed of light in vacuum and m_r and m_i are the real and imaginary parts of the refractive index, respectively. Substituting Eq. (5.4.9c) into Eq. (5.4.14) and using Eq. (5.4.10) lead to

$$N_r^2 - \tilde{N}_i^2 = m_r^2 - m_i^2, \quad N_r\tilde{N}_i \cos\theta_t = m_r m_i. \tag{5.4.15}$$

Let $N_i = \tilde{N}_i \cos\theta_t$. Then, from Eqs. (5.4.13) and (5.4.15), we obtain

$$N_r = \frac{\sqrt{2}}{2} \left\{ m_r^2 - m_i^2 + \sin^2\theta_i + \left[(m_r^2 - m_i^2 - \sin^2\theta_i)^2 + 4m_r^2 m_i^2 \right]^{1/2} \right\}^{1/2}, \tag{5.4.16a}$$

$$N_i = \frac{m_r m_i}{N_r}. \tag{5.4.16b}$$

These two parameters are referred to as the adjusted real and imaginary refractive indices.

After determining N_r and N_i, the refracted wave given in Eq. (5.4.9c) can be rewritten in the form

$$\mathbf{E}_t(\mathbf{r}, t) = \mathbf{A}_t \exp(-kN_i l_a) \exp\left[i(kN_r\mathbf{e}^t \cdot \mathbf{r} - \omega t) \right], \tag{5.4.17}$$

where $l_a = (\mathbf{e}^\alpha \cdot \mathbf{r}) / \cos \theta_t$ is the distance of the propagation of the refracted wave along the direction $\mathbf{e}^t$. It is clear that the direction of the phase propagation for the inhomogeneous wave inside the medium is determined by N_r via Snell's law, whereas the attenuation of the wave amplitude during the wave propagation is determined by N_i. Consequently, the refracted wave can be traced precisely. The Fresnel reflection and refraction coefficients defined in Eq. (5.3.23) in terms of the adjusted real and imaginary refractive indices can then be written as follows:

$$R_l = \frac{N_r \cos \theta_i - \cos \theta_t}{N_r \cos \theta_i + \cos \theta_t}, \quad T_l = \frac{2 \cos \theta_i}{N_r \cos \theta_i + \cos \theta_t}, \tag{5.4.18a}$$

$$R_r = \frac{\cos \theta_i - N_r \cos \theta_t}{\cos \theta_i + N_r \cos \theta_t}, \quad T_r = \frac{2 \cos \theta_i}{\cos \theta_i + N_r \cos \theta_t}, \tag{5.4.18b}$$

where the subscripts l and r denote the parallel and perpendicular polarized components, respectively.

5.4.1.4   MONTE CARLO METHOD FOR RAY TRACING

Use of the Monte Carlo method in connection with geometric ray-tracing was first developed by Wendling *et al.* (1979) for hexagonal ice columns and plates. Takano and Liou (1995) further innovated a hit-and-miss Monte Carlo method to trace photons in complex ice crystals, including the contributions of absorption and polarization.

Let a bundle of parallel rays, representing a flow of photons, be incident on a crystal from a direction denoted by a set of two angles with respect to the crystal's principal axis. Consider a plane normal to this bundle of incident rays and the geometric shadow of a crystal projected onto this plane. Further, let a rectangle (defined by X and Y) enclose this geometric shadow such that the center of this rectangle coincides with the center of the crystal. One of the sides, X, is parallel to the geometric shadow of the crystal's principal axis. A point (x_i, y_i), is selected inside this rectangle using random numbers, RN, whose range is from 0 to 1 such that

$$x_i = X \left(\text{RN} - \frac{1}{2} \right), \tag{5.4.19a}$$

$$y_i = Y \left(\text{RN} - \frac{1}{2} \right). \tag{5.4.19b}$$

In this manner, x_i is from $-X/2$ to $X/2$, whereas y_i is from $-Y/2$ to $Y/2$. If the point is inside the geometric shadow, it is regarded as an incident point on the crystal. Otherwise, it is disregarded. If there are more than two crystal planes for a photon, the point closer to the light source is regarded as the incident point. The coordinates of an incident point (x_i, y_i) can be transformed into the coordinates (x, y, z) with respect to the body-framed coordinate system using the method described by Takano and Asano

(1983) for efficient geometric ray-tracing procedures. Once the incident coordinates are determined, the photons are traced with a hit-and-miss Monte Carlo method. The Fresnel reflection coefficients, R_l and R_r, are first calculated and compared with a random number, RN. If $(|R_l|^2 + |R_r|^2)/2$ is greater than RN, the photon is reflected. Otherwise, it is transmitted. When a photon traverses a particle, it can be absorbed. One can account for absorption by means of a stochastic procedure. When a photon enters a crystal, an absorption path length, l_a, is generated with a random number such that

$$\text{RN} = \exp(-2kN_i l_a), \quad \text{i.e., } l_a = -\ln\left(\frac{\text{RN}}{2kN_i}\right). \tag{5.4.20}$$

The random number represents the probability of the transmission of a photon. The absorption path length, l_a, denotes the distance traversed by a photon in the crystal before the photon is absorbed. An actual path length, l, between an incident point and the next internal incident point can then be calculated on the basis of Snell's law and the specific ice-crystal geometry. The transmission is then given by $T = \exp(-2kN_i l)$. If $T \leq \text{RN} \leq 1$, then the photons associated with these RNs are absorbed. Similarly, if l is greater than l_a, then the photon is absorbed. Otherwise, it is transmitted without absorption. This procedure is repeated whenever photons travel inside the crystal.

After a photon is transmitted out of the crystal or reflected externally, it can reenter the crystal depending on the crystal's shape. In this case, a new incident direction can be calculated using the direction cosine of the scattered beam. The new incident coordinates can also be determined from the new incident direction and the coordinates of an emergent point of the photon on the crystal surface. The foregoing procedure is repeated until the photon escapes from the crystal. When a photon reenters the crystal, the scattering angle and the scattering matrix are computed with respect to the original incident direction. In the conventional method, the number of scattered photons per unit solid angle, $2\pi \sin\theta\,\Delta\theta$, is counted as the phase function. The single-scattering albedo is obtained from the ratio of the number of scattered photons to the number of incident photons. The Monte Carlo method allows us to treat complicated ice crystals effectively and can be employed in connection with the improved geometric ray-tracing approach.

The surfaces of ice crystals may not be exactly smooth, particularly if they undergo collision processes. Also, a careful examination of some polycrystalline ice crystals reveals rough structures on the surfaces. Halo and arc patterns that are absent from some cirrus clouds could be caused by deviations of the ice-crystal surfaces from defined hexagonal structures. Incorporation of some aspects of ice crystal surface roughness in geometric ray-tracing can be made by following the idea developed by Cox and Munk (1954) for wavy sea surfaces. A rough surface may be thought of as consisting of a number of small facets that are locally planar and randomly tilted from the flat surface. We may use a two-dimensional Gaussian probability function

to define the surface tilt as follows:

$$p(z_x, z_y) = \frac{1}{\pi \sigma^2} \exp\left(-\frac{z_x^2 + z_y^2}{\sigma^2}\right), \qquad (5.4.21a)$$

with

$$z_x = \frac{\partial z}{\partial x} = \tan \theta \cos \phi, \qquad (5.4.21b)$$

$$z_y = \frac{\partial z}{\partial y} = \tan \theta \sin \phi, \qquad (5.4.21c)$$

where z_x and z_y are the slopes defined for a facet of rough surface along two orthogonal directions, θ and ϕ are the local polar angles defining the position of the tilt of the surface facet, and σ is a parameter controlling the degree of roughness. In general, surface roughness of ice particles has the effect of smoothing out the scattering maxima that occur in the phase function.

5.4.2   Introduction to the Finite-Difference Time Domain Method

To circumvent the inherent shortcomings in the geometric optics approach due to the requirement of localization of light rays, we introduce a specific numerical method referred to as the *finite-difference time domain* (FDTD) method for light scattering by small ice crystals. Details of this method will be elaborated on in the discussion of light scattering by nonspherical aerosols in Section 5.5.1. For the continuity of this presentation, however, we shall address the physical fundamentals of the methodology.

The FDTD technique is a direct implementation of the Maxwell curl equations to solve the temporal variation of electromagnetic waves within a finite space containing the scatterer. The three-dimensional scatterer must be discretized into a number of suitably selected rectangular cells, referred to as grid meshes, in which the optical properties are defined. Discretizations are subsequently carried out for the Maxwell curl equations by using the finite difference approximation in both time and space. The propagation and scattering of the excited wave in the time domain can be simulated from the discretized equations by a method of time-marching iterations.

In numerical computations, scattering of the electromagnetic wave by a particle must be confined to finite space. In the application of the FDTD technique, therefore, it is necessary to impose artificial boundaries so that the simulated field within the truncated region would be the same as that in the unbounded case. Implementation of an efficient absorbing boundary condition to suppress spurious reflections is an important aspect of the FDTD method associated with numerical stability and computer time and memory requirements.

The solution of the finite difference analog of the Maxwell curl equations is in the time domain. To obtain the frequency response of the scattering particle, we require an appropriate transformation. The discrete Fourier transform technique can be employed to obtain the frequency spectrum of the time-dependent signals if a Gaussian pulse

is used as an initial excitation. Correct selection of the pulse is required to avoid numerical aliasing and dispersion.

Finally, mapping of the near-field results to the far field must be performed to derive the scattering and polarization properties of the particle. A surface integration or a volume integration technique can be employed to obtain the far-field solution. Numerical calculations of the FDTD method pose some fundamental problems, including the staircasing effect in approximating the particle shape and the absorbing boundary condition used to truncate the computational domain. In Section 5.5.1, we will demonstrate that the FDTD approach can be applied to size parameters smaller than about 20 with adequate accuracy.

5.4.3   Scattering Phase Matrix for Nonspherical Ice Particles

Consider a particle of arbitrary shape and size. The scattered electric field at a distance r from the particle must be related to the two components of the incident electric field (E_l^i, E_r^i). In the far field, the two-by-two amplitude matrix that transforms the incident electric vector into the scattered electric vector may be written in the form

$$\begin{bmatrix} E_l^s \\ E_r^s \end{bmatrix} = \frac{\exp(-ikr + ikz)}{ikr} \begin{bmatrix} S_2 & S_3 \\ S_4 & S_1 \end{bmatrix} \begin{bmatrix} E_l^i \\ E_r^i \end{bmatrix}, \tag{5.4.22}$$

where z is the vertical direction in the Cartesian coordinates and S_j $(j = 1, 2, 3, 4)$ are the amplitude functions. For spherical particles, $S_3 = S_4 = 0$, as shown in Eq. (5.2.83).

In terms of the Stokes parameters defined in Eq. (5.2.103), we find

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \frac{\mathbf{F}}{k^2 r^2} \begin{bmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{bmatrix}, \tag{5.4.23}$$

where the subscript 0 denotes the incident beam and the transformation matrix is given by

$$\mathbf{F} = \begin{bmatrix} \frac{1}{2}(M_2 + M_3 + M_4 + M_1) & \frac{1}{2}(M_2 - M_3 + M_4 - M_1) & S_{23} + S_{41} & -D_{23} - D_{41} \\ \frac{1}{2}(M_2 + M_3 - M_4 - M_1) & \frac{1}{2}(M_2 - M_3 - M_4 + M_1) & S_{23} - S_{41} & -D_{23} + D_{41} \\ S_{24} + S_{31} & S_{24} - S_{31} & S_{21} + S_{34} & -D_{21} + D_{34} \\ D_{24} + D_{31} & D_{24} - D_{31} & D_{21} + D_{34} & S_{21} - S_{34} \end{bmatrix}. \tag{5.4.24}$$

Expressions for the matrix elements in terms of the electric fields can be derived from the definition of the Stokes parameters and are given by

$$M_k = |S_k|^2, \tag{5.4.25a}$$

$$S_{kj} = S_{jk} = (S_j S_k^* + S_k S_j^*)/2, \tag{5.4.25b}$$

$$-D_{kj} = D_{jk} = (S_j S_k^* - S_k S_j^*)i/2, \quad j, k = 1, 2, 3, 4. \tag{5.4.25c}$$

The preceding matrix elements are real numbers. In radiative transfer, it is conventional to define the scattering phase matrix, **P**, such that its first element is normalized to unity as follows:

$$\int_0^{2\pi} \int_0^\pi \frac{P_{11}(\theta)}{4\pi} \sin\theta \, d\theta \, d\phi = 1, \tag{5.4.26}$$

where θ and ϕ denote the scattering and azimuthal angles, respectively.

Next, we define the scattering cross section, which represents the amount of incident flux that is removed from the original direction as a result of a single-scattering event such that this flux is distributed isotropically throughout the area of a sphere whose radius is R and whose center is the scatterer. The scattering cross section is related to the first element of the scattering phase matrix in the form

$$\sigma_s = \frac{1}{k^2} \int_0^{2\pi} \int_0^\pi \left(\frac{1}{2} \sum_{k=1}^4 M_k \right) \sin\theta \, d\theta \, d\phi. \tag{5.4.27}$$

Using Eq. (5.4.27), the scattering phase matrix may be defined in terms of the transformation matrix as follows:

$$\frac{\mathbf{P}}{4\pi} = \frac{1}{\sigma_s k^2} \mathbf{F}. \tag{5.4.28}$$

The Stokes parameters can then be expressed in the form

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \Omega_{\text{eff}} \frac{\mathbf{P}}{4\pi} \begin{bmatrix} I_0 \\ Q_0 \\ U_0 \\ V_0 \end{bmatrix}, \tag{5.4.29}$$

where $\Omega_{\text{eff}} = \sigma_s/r^2$, denoting the effective solid angle associated with scattering. If no assumption is made about the shape and position of the scatterer, the scattering phase matrix consists of 16 nonzero elements:

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix}. \tag{5.4.30}$$

The preceding discussion is concerned with the scattering of light by a nonspherical particle. For a group of nonspherical particles, its scattering property is determined by the orientation and size of individual particles with respect to the incident light beam. For continuity of the present discussion, we shall consider a sample of nonspherical particles of the same size randomly oriented in space. The scattering phase matrix can then be expressed by

$$\mathbf{P}(\theta) = \frac{1}{2\pi \sigma_s} \int_0^{2\pi} \int_0^{\pi/2} \mathbf{P}'(\alpha', \gamma') \sigma_s'(\alpha', \gamma') \sin\alpha' \, d\alpha' d\gamma', \tag{5.4.31a}$$

where α' and γ' are the orientation angles of a nonspherical particle with respect to the incident light beam (see Fig. 6.17 and Section 6.7.1 for further discussion), and

$\mathbf{P}'$ denotes the scattering phase matrix for a single particle. In this case, the scattering phase matrix is a function of the scattering angle only. Further, the scattering cross section defined in Eq. (5.4.27) for randomly oriented nonspherical particles can be written in the form

$$\sigma_s = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\pi/2} \sigma_s'(\alpha', \gamma') \sin \alpha' d\alpha' d\gamma'. \tag{5.4.31b}$$

Again, σ_s' denotes the scattering cross section for a single particle. For a symmetric hexagonal crystal, the azimuthal orientation angle γ' ranges from 0 to $\pi/6$ and an analytic expression can be obtained in this case (Exercise 5.14).

If the scatterers are randomly oriented in space such that every scatterer has a plane of symmetry, the law of reciprocity can be applied (Perrin, 1942; van de Hulst, 1957). Thus, we can reverse the directions of the incident and scattered polarized beams, and achieve the same results. It follows that the amplitude functions (S_3, S_4) in Eq. (5.4.23a) must be equivalent to $(-S_3, -S_4)$. Consequently, the following six relationships between the phase matrix elements are valid: $P_{12} = P_{21}$, $P_{13} = -P_{31}$, $P_{14} = P_{41}$, $P_{23} = -P_{32}$, $P_{24} = P_{42}$, and $P_{34} = -P_{43}$ (Exercise 5.15). Moreover, consider an incident light beam described by negative ellipticity and orientation angles (see Section 6.6.1) and whose Stokes parameters are $(I_0, Q_0, -U_0, -V_0)$. The scattered beam from an isotropic medium composed of randomly oriented scatterers must have the same form for the Stokes parameters: $(I, Q, -U, -V)$. The (I, Q) components by definition are invariant to the change in the incident Stokes parameters from (U, V) to $(-U, -V)$. Using Eq. (5.4.29), four relationships can be derived such that $P_{13} = P_{14} = P_{23} = P_{24} = 0$, and $P_{31} = P_{32} = P_{41} = P_{42} = 0$. It follows that the scattering phase matrix reduces to six independent elements in the form

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & 0 & 0 \\ P_{12} & P_{22} & 0 & 0 \\ 0 & 0 & P_{33} & P_{34} \\ 0 & 0 & -P_{34} & P_{44} \end{bmatrix}. \tag{5.4.32}$$

For spherical scatterers, $S_3 = S_4 = 0$. Thus, $P_{22} = P_{11}$ and $P_{44} = P_{33}$. Consequently, there are only four independent scattering phase matrix elements (Section 5.2.4).

Consider now a spectrum of nonspherical particles randomly oriented in space with a size distribution given by $n(L)$, where L is the major axis of a nonspherical particle. In the case of ice particles or aerosols, we use a parameter referred to as the aspect ratio $L/2a$, where $2a$ denotes the width, to define the particle size. We may employ a mean effective size in a manner defined in Eq. (5.1.1) in association with light scattering calculations. The scattering phase matrix associated with a suitable volume, which contains a spectrum of nonspherical particles defined by a size distribution $n(L)$, can be obtained from

$$\mathbf{P}(\theta) = \int_{L_1}^{L_2} \mathbf{P}(\theta, L)\sigma_s(L)n(L)dL \bigg/ \int_{L_1}^{L_2} \sigma_s(L)n(L)dL, \tag{5.4.33a}$$

where L_1 and L_2 are the lower and upper limits of the particle length, and the scattering cross section for a spectrum of nonspherical particles of different sizes and shapes is

given by

$$\sigma_s = \int_{L_1}^{L_2} \sigma_s(L)n(L)\,dL/N, \qquad (5.4.33b)$$

where

$$N = \int_{L_1}^{L_2} n(L)\,dL \qquad (5.4.33c)$$

is the total number of nonspherical particles. A similar definition can be derived for the extinction cross section. It follows that a single-scattering albedo for a spectrum of nonspherical particles is given by $\tilde{\omega} = \sigma_s/\sigma_e$.

As stated previously, the condition under which the scattering phase matrix is composed of six independent elements is that the nonspherical particles are randomly oriented in space in such a manner that every one of them has a plane of symmetry based on which the law of reciprocity can be applied: that the incident and scattered beams are reversible. In Section 5.1, we illustrated examples of typical ice-crystal sizes and shapes. Assuming that they are randomly oriented in space, it is very difficult, if not impossible, to prove that every one of these ice crystals has a plane of symmetry that would obey the law of reciprocity. Nevertheless, numerical calculations for light scattering by ice crystals carried out by Takano and Liou (1989a, 1995) based on the geometric ray-tracing approach show that the scattering phase matrix elements P_{13}, P_{14}, P_{23}, P_{24}, P_{31}, P_{32}, P_{41}, and P_{42} are practically zeros and that $P_{12} = P_{21}$ and $P_{34} = -P_{43}$. Light scattering and radiative transfer dealing with the possibility of ice crystals oriented horizontally in the atmosphere are discussed further in Section 6.7.1. In the following, we present the essence of the unified theory for light scattering by ice crystals, comparison to available measurements, and some representative single-scattering results.

5.4.4   Presentation of a Unified Theory for Light Scattering by Ice Crystals

5.4.4.1   THE ESSENCE OF THE UNIFIED THEORY

In recognition of the inherent limitations of each methodology for light scattering calculations, it is unlikely that one specific method can be satisfactorily employed to resolve the intricate scattering problems involving nonspherical ice crystals of all size parameters and shapes. However, by unifying the geometric optics and FDTD methods described earlier, calculations of light scattering and absorption by ice crystals covering all sizes and shapes that commonly occur in the atmosphere can be performed with adequate precision. This approach is referred to as the unified theory for light scattering by ice crystals (Liou *et al.*, 2000). In Fig. 5.19, we demonstrate this theory in terms of the extinction efficiency as a function of size parameter for randomly oriented columns of a uniform size. The length-to-width ratio for the column is 6 and the incident wavelength used is 0.63 μm. The FDTD results are presented for size parameters from 1 to 30 for comparison purposes. The results computed from the improved geometric optics method cover size parameters from 1 to 1000. Significant deviations between these two methods occur when the size parameter is
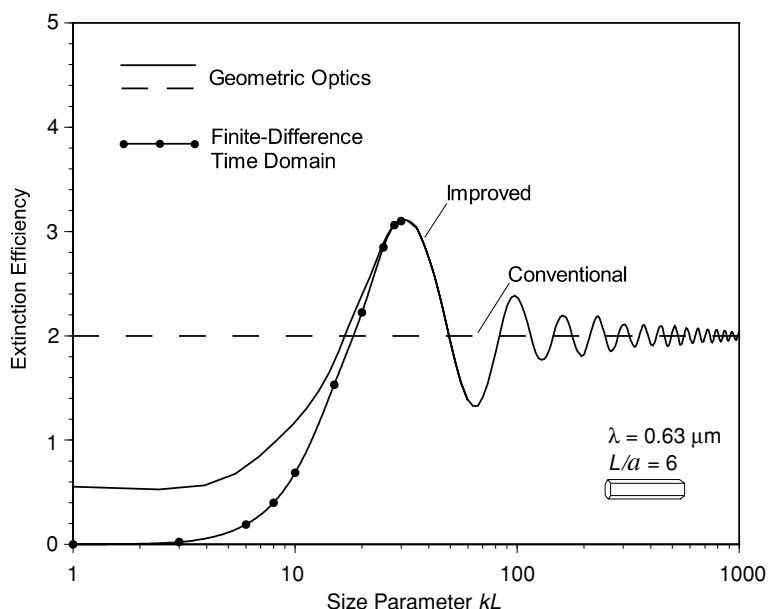
Figure 5.19   Presentation of a unified theory for light scattering by ice crystals using the extinction efficiency as a function of size parameter as an example. The solid line represents results computed from the improved geometric optics method, while the dots are results computed from the finite-difference time domain method. The dashed line represents the conventional extinction efficiency of 2 based on the principle of geometric optics. The presentations use a wavelength of 0.63 μm and randomly oriented ice columns of a uniform size identified in the figure.

smaller than about 15. This comparison highlights the increasing inapplicability of the localization principle inherent in geometric ray-tracing for the calculation of the electric field at the surface of a scatterer with a small size parameter. Based on the optical theorem noted previously, the extinction efficiency is equal to 2 regardless of the size parameter in the conventional geometric optics method. A series of systematic comparisons has been carried out for the phase function, single-scattering albedo, and extinction cross section computed by the improved geometric optics and FDTD methods for solid and hollow columns, plates, bullet rosettes, and aggregates. The improved geometric optics method was shown to produce acceptable accuracy for the single-scattering parameters for these ice crystal shapes with size parameters on the order of 15–20. Also, in these calculations, the peaks in the 22° and 46° halos generated by the geometric ray-tracing were found to be substantially smaller when the size parameter is smaller than about 60. Some of the cirrus clouds that do not produce the well-known halos and arcs could involve small ice crystals and/or the deviation of large ice crystals from the hexagonal geometries caused by the growth process.

The unified theory for light scattering by ice crystals is a powerful tool for the calculation of the phase function and other optical parameters. Figure 5.20 displays various commonly observed ice-crystal habits generated by computer and their associated phase functions. Eleven ice-crystal types are shown, including solid and
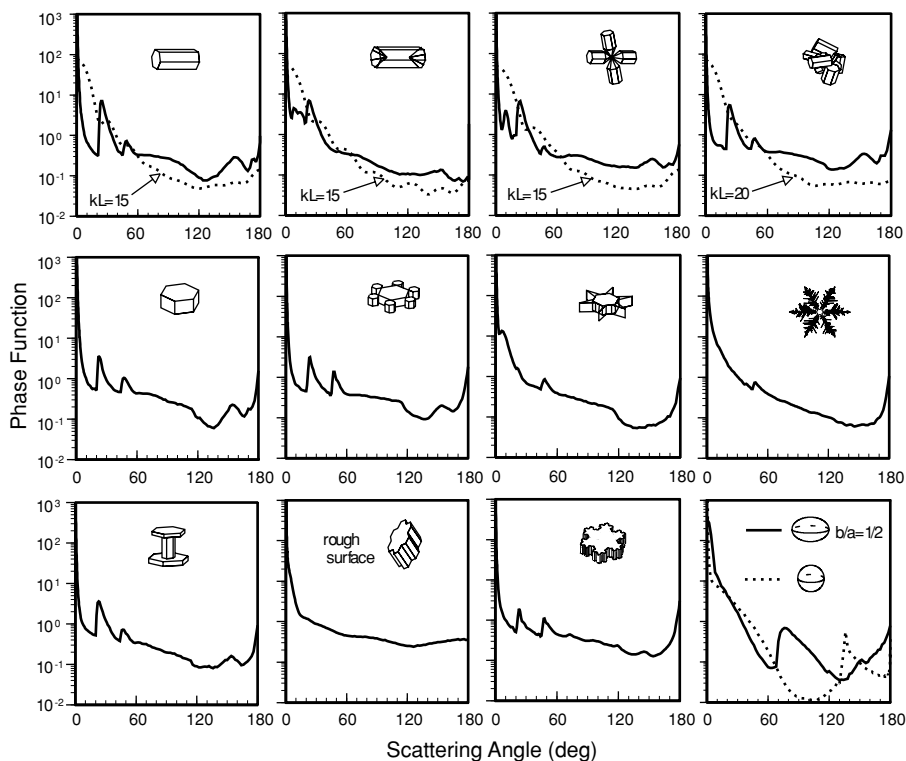
Figure 5.20   Eleven ice-crystal habits commonly occurring in cirrus clouds generated by computer program, along with their associated phase function patterns for the 0.63 μm wavelength. The solid lines are results computed from the geometric ray-tracing method; the dashed lines in the upper panel are results generated from the finite-difference time domain technique for which the size parameters are indicated in the diagrams. Also shown for comparison are phase functions for a spheroid and a sphere (after Liou *et al.,* 2000).

hollow columns, single and double plates and a plate with attachments, a bullet rosette with four branches, an aggregate composed of eight columns, a snowflake, a dendrite, and an ice particle with rough surfaces defined by a two-dimensional Gaussian probability function. Also displayed are an ice sphere and an ice spheroid whose phase functions are computed by the exact theory. The phase function for the sphere shows a maximum at the rainbow angle of about 138° and a broad minimum at the side-scattering directions 60–120°, whereas the spheroid displays a maximum pattern in this scattering-angle range. Both phase functions deviate significantly from those of ice particles. The patterns for the snowflake, the dendrite, and the plate with attachments were produced by a fractal shape-generation program. The phase function patterns for the hexagonally based ice crystals, such as solid column, single and double plates, bullet rosette, and column aggregate, all display a strong 22° halo peak (two refractions through a 60° prism angle), a halo peak at 46° (two refractions

through a 90° prism angle), and a maximum at about 160° (one and two internal reflections). The backscattering peak for these ice crystals is produced by external and internal reflections. The bullet also exhibits a peak at about 7°. For plates with small attachments and snowflakes, the 22° and 46° halo peaks are also produced from geometric ray tracing. In the case of the hollow column, the 46°, 160°, and backscattering maxima are absent because of the lack of scattering material. The 22° and 160° maxima vanish for simple and complicated dendrites, but the backscattering is still strong because of internal reflection contributions. The results for small ice crystals computed from the FDTD method are illustrated in the top four panels overlying those computed from geometric ray tracing. Phase function differences between large and small ice crystals are clearly illustrated in the graphs.

5.4.4.2 THEORY VERSUS MEASUREMENT AND REPRESENTATIVE RESULTS

Measurements of the scattering and polarization patterns of ice crystals have been performed in cold chambers (e.g., Sassen and Liou, 1979; Volkovitsky *et al.,* 1980). Desirable ice crystal sizes and shapes, however, are difficult to generate and sustain for a period of time long enough to perform light-scattering experiments. Barkey *et al.* (1999) conducted an experimental light-scattering program using hexagonal ice-like crystals measured in the analog manner so that optical experiments could be performed over a relatively long period of time for complex-shaped particles. The experiment consisted of a polarized laser beam at $\lambda = 0.63$ μm and an array of 36 highly sensitive photodiode detectors arranged between the scattering angles of 2.8° and 177.2° mounted in a linear array on a half-dome, which could be rotated to vary the azimuthal angle. After careful calibration and signal acquisition, this system was first used to measure the phase functions of a glass sphere and a glass fiber configured to scatter light like an infinite cylinder. The experimental results closely match those computed from the Lorenz–Mie theory. Subsequently, sodium fluoride (NaF crystal), which has an index of refraction (1.33) close to that of ice in the visible, was adopted for the light scattering experiment. The NaF crystal could be machined to a desirable shape with a size on the order of a millimeter. It was mounted on top of a small pedestal, and its orientation position was controlled by a rotator. Angular integrations in the experiment could follow the computational procedures that are carried out in theory.

Figure 5.21a shows a comparison between measurements and theory for an aggregate that was assembled from NaF columns with small glass-fiber attachments glued onto small holes. To simulate random orientation, a 1° increment was used for all possible orientation angles. Results derived from the measurements and theory are shown to be in general agreement. Several discrepancies should be noted. Particularly, the experimental results are lower than the theory in backscattering directions, which are dominated by internal reflections. This difference could be caused by absorption by the small glass fibers and glues that connect the columns. Comparison results for a rough-surface plate are shown in Fig. 5.21b. All eight sides were sanded with small scratches evenly distributed across the crystal surface. Between the 25° and 180° scattering angles, the measurements closely follow the theoretical results. For scattering angles less than 20°, however, the experimental results are higher. The
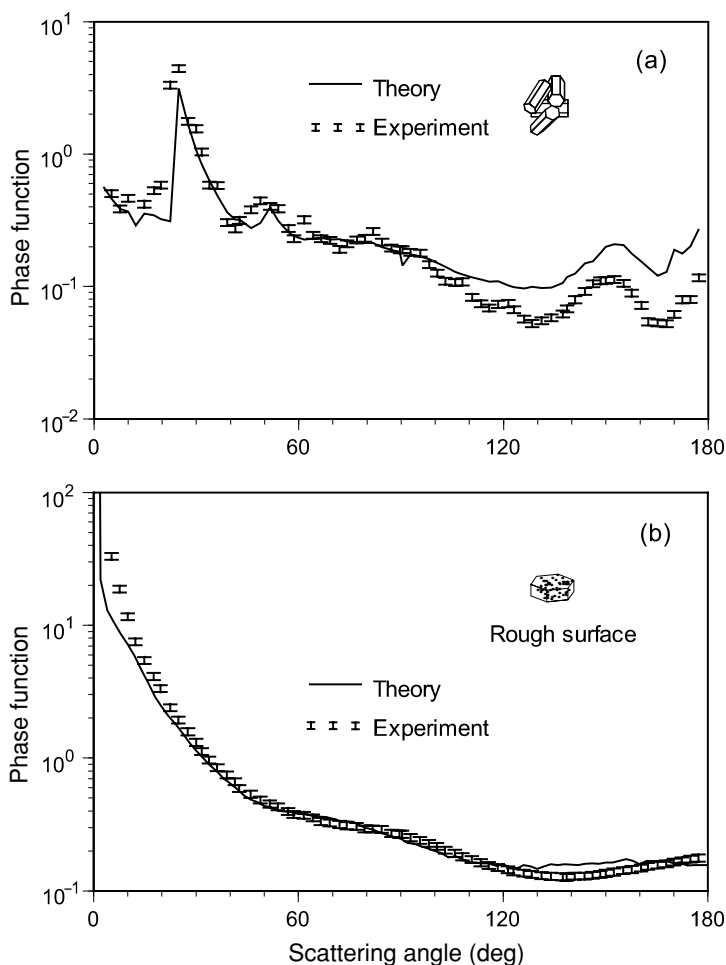
Figure 5.21   Phase functions for (a) randomly oriented aggregate and (b) rough surface plate crystals made from NaF whose index of refraction is 1.33 in the visible. The experiment used a polarized laser beam at 0.63 μm as the light source and controlled the positions of the detector and the crystal by automatic mechanical devices (Barkey *et al.,* 1999). The theoretical results were derived from the geometric-ray-tracing/Monte-Carlo method.

scanning electron photomicrographs reveal features on the roughened crystal surface on the order of 0.5–1 μm. More light could have been scattered through them as compared to the defined cross-sectional area used in diffraction calculations.

The electrodynamic levitation technique has also been used to suspend and grow an individual ice crystal for light-scattering experiments (Bacon *et al.,* 1998). The apparatus consists of an electrodynamic balance with an internally mounted thermal diffusion chamber, a laser beam, a 1024-element linear photodiode array, and two cameras to provide top and side views of the ice crystal. Shown in Fig. 5.22 are
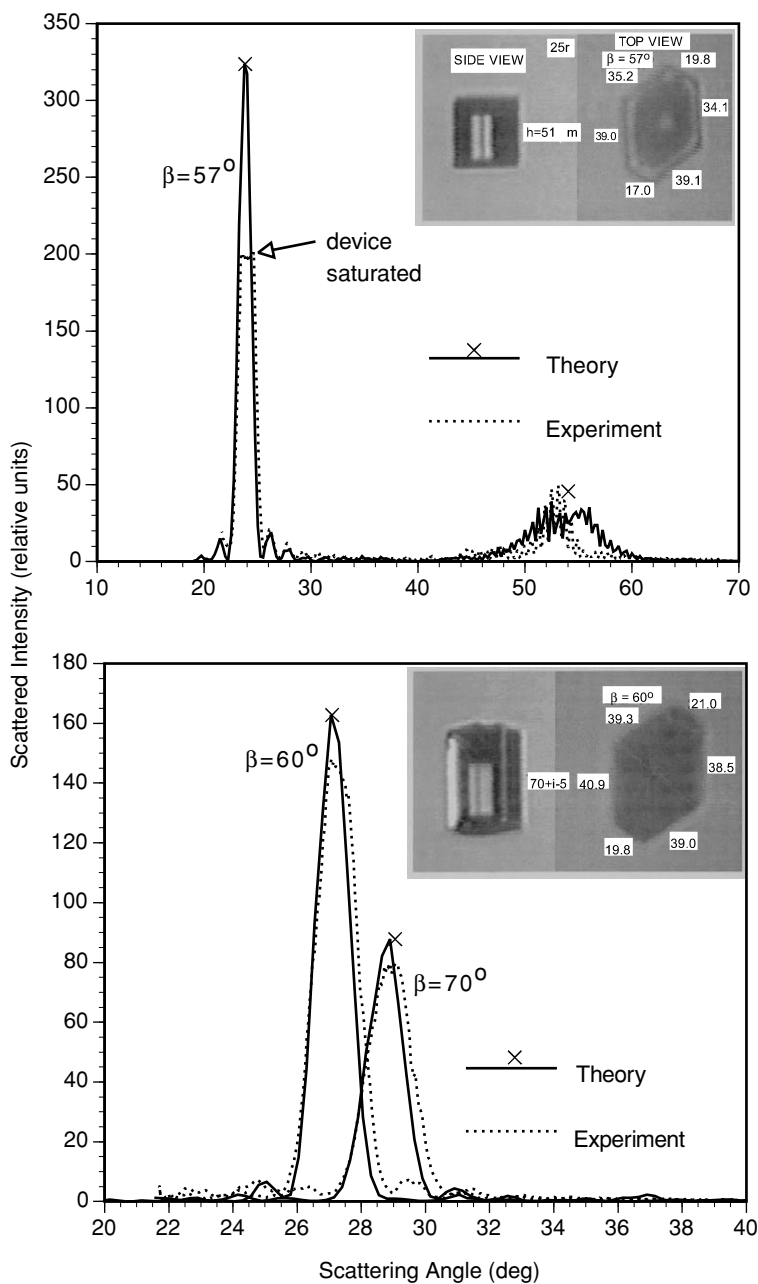
Figure 5.22 Phase function measurements using a laser beam at 0.63 μm and a large number of photodiode detector array for a single ice crystal suspended by the electrodynamic levitation technique (Bacon *et al.*, 1998). The sizes and shapes are determined from the top and side views of two cameras. The angle β denotes the ice-crystal orientation with respect to the incident laser beam. Theoretical results are computed from the conventional ($\times$) and improved geometric ray-tracing methods. Two ice-crystal cases are shown (after Liou *et al.*, 2000).
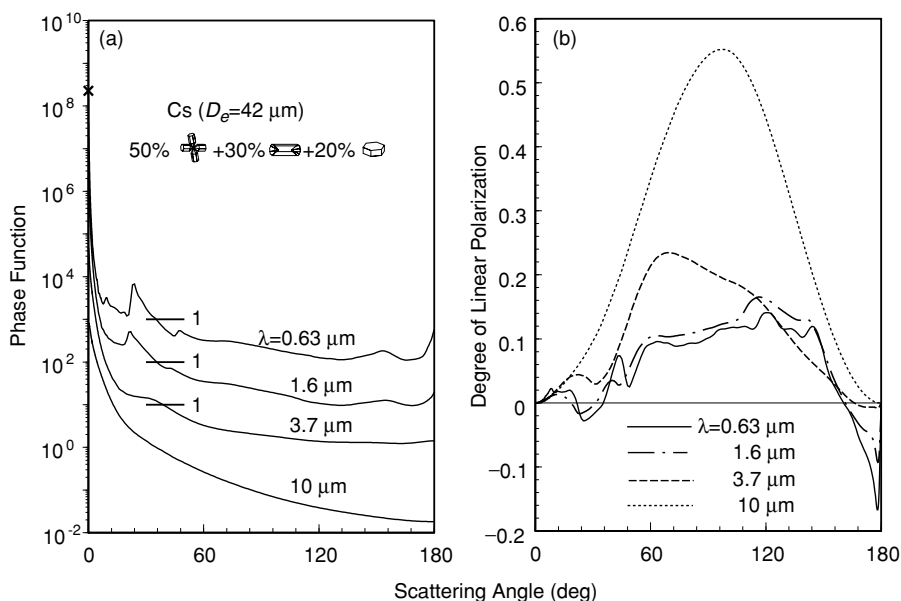
Figure 5.23 (a) Phase function and (b) degree of linear polarization for a typical cirrostratus with a mean effective ice crystal size of 42 μm composed of 50% bullet rosettes/aggregates, 30% hollow columns, and 20% plates, a shape model based on replicator and optical probe measurements. Four remote sensing wavelengths are displayed. For size parameters less than 15, the finite-difference time domain method is employed in the calculations. In the phase function, the vertical scale is applied to the lowest curve, while the upper curves are displayed upward by a factor of 10. The symbol $\times$ denotes the diffraction peak for the 0.63 μm wavelength.

the experimental results for two ice-crystal sizes and shapes defined by the depicted photos. Theoretical results computed from conventional geometric ray-tracing, which does not account for phase interferences, show discrete maxima associated with the halo pattern. The improved geometric optics method generates closely matched patterns, with the exception of some deviations in the 55° scattering angle region in the top diagram. Differences between the theoretical and experimental results can be attributed to the uncertainty of the measurement of the ice-crystal size ($\sim$ 4 μm) and of the computation of the near field based on the geometric ray-tracing approximation.

In Fig. 5.23 we illustrate the phase function and degree of linear polarization for a representative cirrostratus ice crystal size distribution (Section 5.1) having a mean effective ice-crystal size of 42 μm. Based on replicator and optical probe measurements, a cirrus cloud model composed of 50% bullet rosettes/aggregates, 30% hollow columns, and 20% plates is used in this presentation. Four remote sensing wavelengths of 0.63, 1.6, 3.7 and 10 μm are employed in the calculations. As pointed out in Section 5.3.3, these wavelengths are typical of the image channels on meteorological satellites. The scattering and polarization results for other remote-sensing wavelengths, 0.86 μm and 10.9 and 11 μm, are similar to those for 0.63 and 10 μm,

respectively. In the phase function, the 0.63 and 1.6 μm wavelengths display a peak at the 22° halo position because of the basic hexagonal structure of the rosettes, columns, and plates used in the calculations. Because of ice absorption, all the features associated with hexagonal ice crystals vanish at 3.7 and 10 μm, except for the forward diffraction peak. The linear polarization patterns for single scattering show a number of maxima features at the 0.63 and 1.6 μm wavelengths: a slight negative polarization at the 22° peak, a positive polarization at the 46° peak, and a strong negative polarization close to the backscattering angle. From about 60° to 140°, about 10% polarization is observed. The case of 10 μm exhibits a maximum at about 90° produced principally by external reflection and limited internal reflections of the light beam.

The spectral single-scattering parameters for ice crystals in terms of the asymmetry factor, single-scattering albedo, and extinction coefficient covering the spectral intervals of 0.2–5 μm and 8–13 μm are presented in Fig. 5.24 using the same cirrus cloud model shown in Fig. 5.23. Three mean effective ice-crystal sizes ranging from 10 μm (contrail cirrus), to 42 μm (typical cirrostratus), to 124 μm (cirrus uncinus), were selected for this illustration. The extinction coefficient is normalized in reference to the value at 0.5 μm. In the visible and near IR wavelengths, the average size parameter is sufficiently large that the extinction efficiency is approximately equal to 2 on the basis of the optical theorem, except in the vicinity of about 3 μm where a significant dispersion of the real part of the refractive index m_r for ice occurs (see also Fig. D.1). The single-scattering albedo pattern mimics the variability of the imaginary refractive index for ice (Liou, 1992, Fig. 5.1) with a large minimum located at about 3 μm. The single-scattering albedo decreases as the mean effective size increases, with an exception in the vicinity of about 2.85 μm, referred to as the *Christiansen effect*. This effect occurs when the real part of the refractive index approaches 1, while the corresponding imaginary part is substantially large, resulting in the domination of absorption. At the wavelength of about 10.9 μm, m_r for ice is close to 1.08, which is the smallest in the window region. However, m_i increases in the window wavelengths, the combination of which leads to a smaller value for extinction for the smallest size parameter of 10 μm. We also notice that the single-scattering albedo displays a significant drop at 10.9 μm. Because of this unique property of ice crystals, inference of their size appears feasible by using the window wavelengths that exist in present and future satellite systems (see Section 7.4.5, the subsection on information content in the IR line spectrum, for additional discussion).

5.5 Light Scattering by Nonspherical Aerosols

In Section 5.1, we showed that atmospheric aerosols exhibit various shapes with a typical size less than about 1 μm. The Lorenz–Mie theory of light scattering by spheres presented in Section 5.2 cannot be applied to nonspherical aerosol particles, nor can the geometric optics approach introduced in Section 5.3 be used for scattering and absorption calculations for aerosols because of their small size parameters. Many
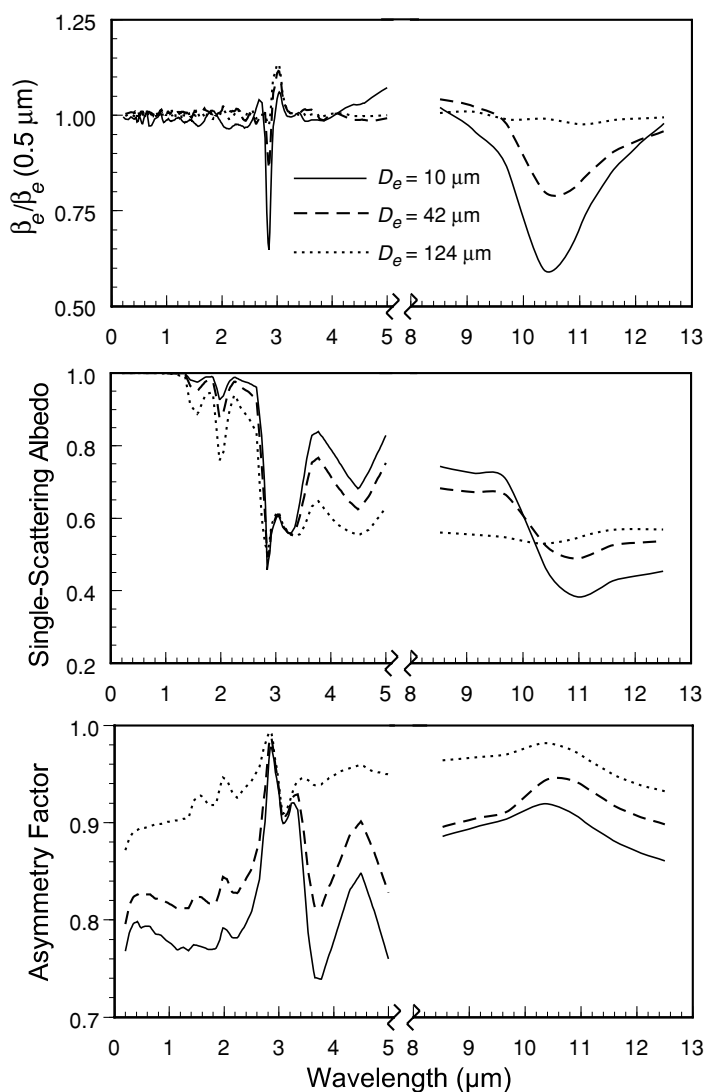
Figure 5.24   Normalized extinction coefficient (in reference to the value at 0.5 μm), single-scattering albedo, and asymmetry factor as functions of wavelength for three ice-crystal size distributions with mean effective sizes of 10, 42, and 124 μm and the shape model defined in Fig. 5.23. The extinction coefficient values at $\lambda = 0.5\,\mu$m are 0.3036, 0.2082, and 1.979 km^{-1} for $D_e = 10$, 42, and 124 μm, respectively. The spectral intervals covered are from 0.2 to 5 μm in the solar spectrum, and from 8 to 13 μm in the thermal infrared window.

approaches have been developed for the calculation of light scattering by nonspherical particles because of the need for precise scattering information in optics, geophysics, remote sensing, astrophysics, engineering, medicine, and biology. In the following sections, we introduce two contemporary methods that have been developed for the calculation of light scattering and absorption by nonspherical aerosols.

5.5.1   Finite-Difference Time Domain Method

The finite-difference time domain (FDTD) technique has been shown to be an efficient computational method of solving for the interaction of electromagnetic waves with scatterers, particularly those with complicated geometry and inhomogeneous composition. In this method, the space containing a scattering particle is discretized by using a grid mesh, and the existence of the particle is represented by assigning suitable electromagnetic constants in terms of permittivity, permeability, and conductivity to the grid points (Fig. 5.25). Because it is not necessary to impose the electromagnetic boundary conditions at the particle surface, the FDTD approach with appropriate and minor modifications can be applied to the solution of light scattering by various small nonspherical and inhomogeneous particles, such as irregular ice crystals and aerosols with inclusions.

The FDTD method was developed and pioneered by Yee (1966), but it did not receive significant recognition until high-quality absorbing boundary conditions were developed in the 1980s. Through the persistent efforts of a number of electrical engineers and computational physicists (e.g., Taflove, 1995), several advantages of the FDTD method have now become widely recognized. In recent years, the FDTD technique has been used to solve for the interaction between targets and electromagnetic
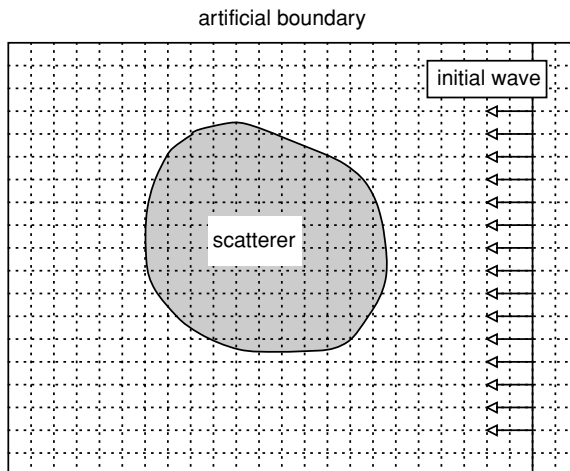

Figure 5.25 A conceptual diagram for the computation of the near field for the scattering of a nonspherical particle illuminated by plane waves by means of the finite-difference time domain (FDTD) method.

waves involving such problems as antenna scattering, numerical modeling of microstrip structures, and electromagnetic absorption by human tissues. Application of this method to the solution of the scattering and polarization properties of atmospheric nonspherical particles has also been carried out by Yang and Liou (1996b, 2000).

As described in Section 5.4.2, the FDTD technique uses Maxwell's time-dependent curl equations, denoted in Eqs. (5.2.10) and (5.2.11), to solve for the temporal variation of electromagnetic waves within a finite space that contains the scattering object. In this case, we write

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\varepsilon(\mathbf{r})}{c} \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t}, \qquad (5.5.1a)$$

$$\nabla \times \mathbf{E}(\mathbf{r}, t) = -\frac{1}{c} \frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t}, \qquad (5.5.1b)$$

where ε is the permittivity of the dielectric medium, usually a complex variable, and c is the speed of light in vacuum. In Eq. (5.5.1b), the permeability, μ, has been assumed to be unity because cloud and aerosol particles in the atmosphere and many other scattering targets are mostly nonferromagnetic materials. We should select a harmonic time-dependent factor of $\exp(-i\omega t)$ for the electromagnetic wave in the frequency domain so as to have a positive imaginary part of the refractive index. Thus, we have

$$\varepsilon = \varepsilon_r + i\varepsilon_i, \qquad (5.5.2a)$$

$$\varepsilon_r = m_r^2 - m_i^2, \quad \varepsilon_i = 2m_r m_i, \qquad (5.5.2b)$$

where $i = \sqrt{-1}$ and m_r and m_i are the real and imaginary parts of the refractive index, respectively. When the medium is absorptive, i.e., m_i is nonzero, a complex calculation is required for Eq. (5.5.1a). To circumvent the complex operation, it is desirable to have an equivalent expression for Eq. (5.5.1a) in which the complex permittivity can be avoided. We may introduce an effective real permittivity, ε', and conductivity, σ, and rewrite the first Maxwell curl equation in a source-dependent form as follows:

$$\nabla \times \mathbf{H}(\mathbf{r}, t) = \frac{\varepsilon'}{c} \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} + \frac{4\pi}{c} \mathbf{J}(\mathbf{r}, t), \qquad (5.5.3a)$$

where the second term on the right-hand side of the equation is due to the contribution of the effective electric current given by

$$\mathbf{J}(\mathbf{r}, t) = \sigma \mathbf{E}(\mathbf{r}, t). \qquad (5.5.3b)$$

Transforming Eqs. (5.5.1a) and (5.5.3a) to the equivalent equations in the frequency domain yields

$$\nabla \times \mathbf{H}(\mathbf{r}) = -ik\varepsilon\mathbf{E}(\mathbf{r}), \qquad (5.5.4)$$

$$\nabla \times \mathbf{H}(\mathbf{r}) = -ik(\varepsilon' + i4\pi\sigma/kc)\mathbf{E}(\mathbf{r}), \qquad (5.5.5)$$

where $k = \omega/c$ is the wavenumber of the electromagnetic wave in free space. From a comparison of Eqs. (5.5.4) and (5.5.5), it is clear that the two equations are equivalent if the following conditions hold: $\varepsilon_r = \varepsilon'$ and $\varepsilon_i = 4\pi\sigma/kc$. It follows that Eq. (5.5.1a) can be rewritten as follows:

$$\boldsymbol{\nabla} \times \mathbf{H}(\mathbf{r}, t) = \frac{\varepsilon_r}{c}\left[\frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} + \tau \mathbf{E}(\mathbf{r}, t)\right], \qquad (5.5.6)$$

where $\tau = kc\varepsilon_i/\varepsilon_r$. At this point we can use Eqs. (5.5.1b) and (5.5.6) to construct the finite-difference analog of the Maxwell curl equations, which involves only real calculations. First, we require discretizing the equation in the time domain. To do this, we rewrite Eq. (5.5.6) in the form

$$\frac{\partial[\exp(\tau t)\mathbf{E}(\mathbf{r}, t)]}{\partial t} = \exp(\tau t)\frac{c}{\varepsilon_r}\boldsymbol{\nabla} \times \mathbf{H}(\mathbf{r}, t). \qquad (5.5.7)$$

Integrating Eq. (5.5.7) over the time interval of $[n\Delta t, (n + 1)\Delta t]$, we obtain

$$\exp[\tau(n + 1)\Delta t]\mathbf{E}^{n+1}(\mathbf{r}) - \exp(\tau n \,\Delta t)\mathbf{E}^n(\mathbf{r})$$

$$= \int_{n\Delta t}^{(n+1)\Delta t} \exp(\tau t)\frac{c}{\varepsilon_r}\boldsymbol{\nabla} \times \mathbf{H}(\mathbf{r}, t)dt$$

$$\approx \Delta t \exp[\tau(n + \tfrac{1}{2})\Delta t]\frac{c}{\varepsilon_r}\boldsymbol{\nabla} \times \mathbf{H}^{n+\frac{1}{2}}(\mathbf{r}). \qquad (5.5.8)$$

If follows that

$$\mathbf{E}^{n+1}(\mathbf{r}) = \exp(-\tau\,\Delta t)\mathbf{E}^n(\mathbf{r}) + \exp(-\tau\,\Delta t/2)\frac{c\Delta t}{\varepsilon_r}\boldsymbol{\nabla} \times \mathbf{H}^{n+\frac{1}{2}}(\mathbf{r}). \qquad (5.5.9)$$

Likewise, for the magnetic field we have

$$\mathbf{H}^{n+\frac{1}{2}}(\mathbf{r}) = \mathbf{H}^{n-\frac{1}{2}}(\mathbf{r}) - c\Delta t\boldsymbol{\nabla} \times \mathbf{E}^n(\mathbf{r}). \qquad (5.5.10)$$

In the foregoing equations the superscript n denotes that the associated field is evaluated at the time step $t = n\Delta t$.

The space containing the scattering particle must be discretized by a number of grid cells. We may use cubic cells in the 3D case and evaluate the components of the electric and magnetic fields of a cell at the staggered locations as shown in Fig. 5.26. The advantage of such staggered positions is that the electromagnetic boundary conditions are guaranteed at the interfaces of the cells, so that the tangential components of the E field and the normal components of the H field are continuous at the interfaces. A spatial location in the discretized space is denoted by the indices $(I, J, K) = (I\Delta x, J\Delta y, K\Delta z)$, and any variable as a function of space and time is defined as

$$F^n(I, J, K) = F(I\Delta x, J\Delta y, K\Delta z, n\Delta t), \qquad (5.5.11)$$

in which Δx, Δy, and Δz are the cell dimensions along the x, y, and z axes, respectively. The permittivity must be homogeneous within each cell. For a given cell with
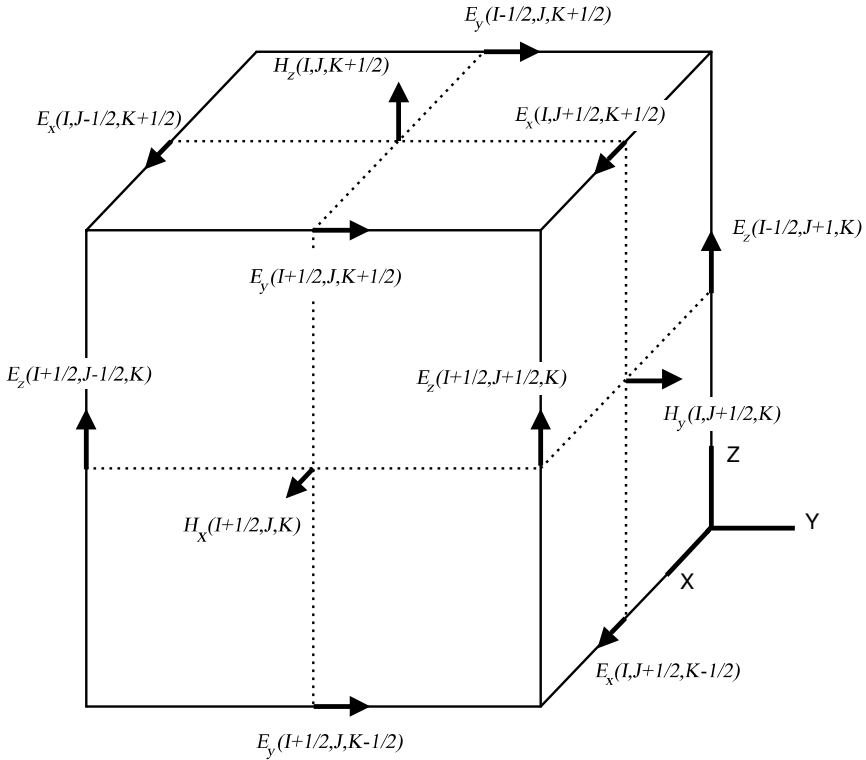
Figure 5.26   Locations of various field components on a cubic cell. The **E** field is tangential to the surface, whereas the **H** field is normal to the surface. The notations I, J, and K are indices assigned to the cell.

its center located at a lattice index (I, J, K), the mean permittivity can be evaluated on the basis of the Maxwell–Garnett rule via

$$\frac{\bar{\varepsilon}(I, J, K) - 1}{\bar{\varepsilon}(I, J, K) + 2} = \frac{1}{\Delta x \Delta y \Delta z} \underset{\text{cell}(I,J,K)}{\int \int \int} \frac{\varepsilon(x, y, z) - 1}{\varepsilon(x, y, z) + 2} \, dxdydz. \quad (5.5.12)$$

Using the mean permittivity produces smaller staircasing errors than those produced by a sharp step-by-step approximation of nonspherical geometry.

The general forms of the finite-difference analogs of Eqs. (5.5.1) and (5.5.6) can then be expressed for each Cartesian component. In practice, the truncation errors of the finite-difference analog of Maxwell's curl equations are of second order in both time and space. The time-marching iteration can be applied directly by employing initial values for the E and H fields. The values of Δx, Δy, Δz, and Δt cannot be specified arbitrarily but must satisfy the Courant–Friedrichs–Levy condition such that

$$c\Delta t \leq \frac{1}{\sqrt{1/\Delta x^2 + 1/\Delta y^2 + 1/\Delta z^2}}. \quad (5.5.13)$$

Moreover, the spatial increments must also be smaller than about 1/20 of the incident wavelength so that the phase variation of the waves is negligible over the distance of the cell dimensions.

The numerical implementation of the FDTD technique requires the imposition of an appropriate absorbing boundary condition, which is critical for the stability of the numerical computations and the reliability of the results. In addition, the white space between the boundary and the scatterer required by a specific boundary condition is an important factor in determining the required computational effort. Many techniques have been developed for the imposition of absorbing boundary conditions in the application of the FDTD technique to electromagnetic scattering problems. A recent advancement in this area is a novel numerical technique, referred to as the *perfectly matched layer* (PML) boundary condition, developed by Berenger (1994). Calculation of the absorption of the outgoing wave by the PML method is based on the absorption by a medium located at the outermost layers in the computational domain backed by a perfectly conducting surface. The conventional technique based on an absorbing medium is to specifically define the wave impedance of the medium so that it matches that of the free space. Such a simple matching approach produces substantial nonzero reflections when a scattered wave impinges on the absorbing medium obliquely. To overcome the disadvantage of the conventional method, the absorbing medium can be selected such that the wave decay due to absorption is imposed on the field components parallel to their boundary layers. To achieve this goal, each Cartesian component of the electromagnetic field is split into two parts as follows:

$$(E_x, E_y, E_z) = \left[(E_{x2} + E_{x3}), (E_{y1} + E_{y3}), (E_{z1} + E_{z2}) \right],$$
$$(H_x, H_y, H_z) = \left[(H_{x2} + H_{x3}), (H_{y1} + H_{y3}), (H_{z1} + H_{z2}) \right], \qquad (5.5.14)$$

where the subscript 1, 2, or 3 denotes the component of the electric (or magnetic) field that is associated with the spatial differential of the magnetic (or electric) field component along the *x, y,* and *z* directions, respectively. With the split field components, discretized scalar equations that govern the propagation of electromagnetic waves can be replaced by a set of 12 equations for numerical computations. Numerical experiments have shown that the spurious reflection produced by the PML boundary conditions is much smaller than that generated by other analytical absorption boundary conditions derived from the wave equation.

The values of the near field computed by the preceding FDTD algorithm are in the time domain. A transformation of the time-dependent field values to their corresponding counterparts in the frequency domain is required to obtain the single-scattering properties. The transformation algorithm depends on the kind of initial wave that is used. We may select an incident pulse (e.g., Gaussian) as the initial excitation in the computation. The width of the pulse can be properly selected to avoid numerical dispersion caused by the finite-difference approximation. In this manner, an incident pulse with respect to a fixed shape and size of the particle can provide the results for a number of size parameters simultaneously. Further, the frequency of the simulated field can be obtained by the discrete Fourier transform. Let *f* be a component of the

field and its value at the time step n be f_n. Then, the time variation of f can be written as

$$f(t) = \sum_{n=0}^{N} f_n \delta(t - n\Delta t), \qquad (5.5.15)$$

where δ is the Dirac delta function and the maximum time step N is chosen such that the field in the time domain is reduced to a small value. The corresponding spectrum in the frequency domain is given by

$$F(k) = \int_{-\infty}^{\infty} \left[\sum_{n=0}^{N} f_n \delta(t - n\Delta t) \right] \exp(ikct) \, dt = \sum_{n=0}^{N} f_n \exp(ikcn\Delta t), \quad (5.5.16)$$

where k is the wavenumber in vacuum. To avoid aliasing and numerical dispersion and to obtain a correct frequency spectrum, one must band the maximum wavenumber or the minimum wavelength for the region within which the frequency response of the scattering is evaluated. In any finite-difference equation, it is required that the wavelength of a simulated wave be larger than the grid size. Finally, the field values in the frequency domain obtained by this procedure must be normalized by the Fourier transform of the incident wave at the center of the grid mesh so that the frequency response of the scattering particle will return to a unit incident harmonic wave.

In order to compute the scattering and absorption quantities, we must transform the near field determined from the FDTD algorithm and the discrete Fourier transform technique to the far field. We may follow an approach that uses a volume integration method for nonconducting scatterers. In the far field (or radiation zone), $k\mathbf{r} \to \infty$, the scattered far field for the electric vector may be written in terms of a volume integral as follows:

$$E_s(\mathbf{r}) = \frac{k^2 \exp(ikr)}{4\pi r} \int\int\int_V [\varepsilon(\mathbf{r}') - 1] \left\{ \mathbf{E}(\mathbf{r}') - \mathbf{n}[\mathbf{n} \cdot \mathbf{E}(\mathbf{r}')] \right\} \exp(-ik\mathbf{n} \cdot \mathbf{r}') d^3\mathbf{r}'.$$

$$(5.5.17)$$

To compute the scattering phase matrix, the scattered field given by Eq. (5.5.17) must be expressed in terms of the amplitude matrix. Because the scattered field is a transverse wave with respect to the scattering direction, it can be decomposed into the components parallel and perpendicular to the scattering plane in the form

$$\mathbf{E}_s(\mathbf{r}) = \alpha E_{s,\alpha}(\mathbf{r}) + \beta E_{s,\beta}(\mathbf{r}), \qquad (5.5.18)$$

where α and β are the unit vectors parallel and perpendicular to the scattering plane, respectively, and satisfy $\mathbf{n} = \beta \times \alpha$. Expressing Eq. (5.5.18) in matrix form yields

$$\begin{pmatrix} E_{s,\alpha}(\mathbf{r}) \\ E_{s,\beta}(\mathbf{r}) \end{pmatrix} = \frac{k^2 \exp(ikr)}{4\pi r} \int\int\int_V [\varepsilon(\mathbf{r}') - 1] \begin{pmatrix} \alpha \cdot \mathbf{E}(\mathbf{r}') \\ \beta \cdot \mathbf{E}(\mathbf{r}') \end{pmatrix} \exp(-ik\mathbf{n} \cdot \mathbf{r}') d^3\mathbf{r}'$$

$$= \frac{\exp(ikr)}{r} \mathbf{S} \begin{pmatrix} E_{o,\alpha} \\ E_{o,\beta} \end{pmatrix}, \qquad (5.5.19)$$

where $\mathbf{S}$ is a 2×2 amplitude scattering matrix and $E_{o,\alpha}$ and $E_{o,\beta}$ are the incident E-field components defined with respect to the scattering plane. In the FDTD method, the incident wave is defined with respect to the incident coordinate system given by $E_{o,x}$ and $E_{o,y}$. Based on the geometry implied by Eq. (5.5.18), we have

$$\begin{pmatrix} E_{o,\alpha} \\ E_{o,\beta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\beta} \cdot \mathbf{x} & -\boldsymbol{\beta} \cdot \mathbf{y} \\ \boldsymbol{\beta} \cdot \mathbf{y} & \boldsymbol{\beta} \cdot \mathbf{x} \end{pmatrix} \begin{pmatrix} E_{o,y} \\ E_{o,x} \end{pmatrix}, \tag{5.5.20}$$

where $\mathbf{x}$ and $\mathbf{y}$ are the unit vectors along the x and y axes, respectively. To obtain the scattering properties of the particle with complete polarization information, we can select two incident cases: (a) $E_{o,x} = 1$ and $E_{o,y} = 0$; and (b) $E_{o,x} = 0$ and $E_{o,y} = 1$. We can then define the following quantities:

$$\begin{pmatrix} F_{\alpha,x} \\ F_{\beta,x} \end{pmatrix} = \frac{k^2}{4\pi} \int\int\int_V [\varepsilon(\mathbf{r}') - 1] \begin{pmatrix} \boldsymbol{\alpha} \cdot \mathbf{E}(\mathbf{r}') \\ \boldsymbol{\beta} \cdot \mathbf{E}(\mathbf{r}') \end{pmatrix} \exp(-ik\mathbf{n} \cdot \mathbf{r}') d^3\mathbf{r}' \bigg|_{E_{o,x}=1,\, E_{o,y}=0},$$
$$\tag{5.5.21a}$$

$$\begin{pmatrix} F_{\alpha,y} \\ F_{\beta,y} \end{pmatrix} = \frac{k^2}{4\pi} \int\int\int_V [\varepsilon(\mathbf{r}') - 1] \begin{pmatrix} \boldsymbol{\alpha} \cdot \mathbf{E}(\mathbf{r}') \\ \boldsymbol{\beta} \cdot \mathbf{E}(\mathbf{r}') \end{pmatrix} \exp(-ik\mathbf{n} \cdot \mathbf{r}') d^3\mathbf{r}' \bigg|_{E_{o,x}=0,\, E_{o,y}=1}.$$
$$\tag{5.5.21b}$$

Using Eqs. (5.5.19)–(5.5.21b) along with some algebraic manipulations, it can be proven that the amplitude matrix is given by

$$\begin{pmatrix} S_2 & S_3 \\ S_4 & S_1 \end{pmatrix} = \begin{pmatrix} F_{\alpha,y} & F_{\alpha,x} \\ F_{\beta,y} & F_{\beta,x} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta} \cdot \mathbf{x} & \boldsymbol{\beta} \cdot \mathbf{y} \\ -\boldsymbol{\beta} \cdot \mathbf{y} & \boldsymbol{\beta} \cdot \mathbf{x} \end{pmatrix}. \tag{5.5.22}$$

The scattering phase matrix can subsequently be defined following the procedure outlined in Section 5.4.3. For nonspherical aerosols oriented randomly in space, the scattering phase matrix has a block-diagonal structure with eight nonzero elements of which only six are independent (Section 5.4.1). Finally, the expressions for extinction and absorption cross sections have been given in Eqs. (5.4.6a) and (5.4.6b), respectively.

Accuracy of the FDTD method can be checked with the exact solutions for infinite circulars and spheres. In general, when the size of the grid cells is on the order of $1/20$ of the incident wavelength, its solutions are in excellent agreement with their corresponding analytical counterparts. Improvement of the accuracy can be achieved by decreasing the ratio of the grid size to the incident wavelength, but only at the expense of increasing the computational effort required. Based on comprehensive numerical experiments, the FDTD method can achieve reliable results for nonspherical aerosol size parameters smaller than about 15–20 (Yang and Liou, 2000). In what follows, we present a number of representative scattering results for nonspherical aerosols.

As illustrated in Section 5.1, aerosols in the atmosphere exhibit a variety of shapes ranging from quasispheres to highly irregular geometries. In addition, aerosols usually appear as a mixed product of different compositions involving dustlike,
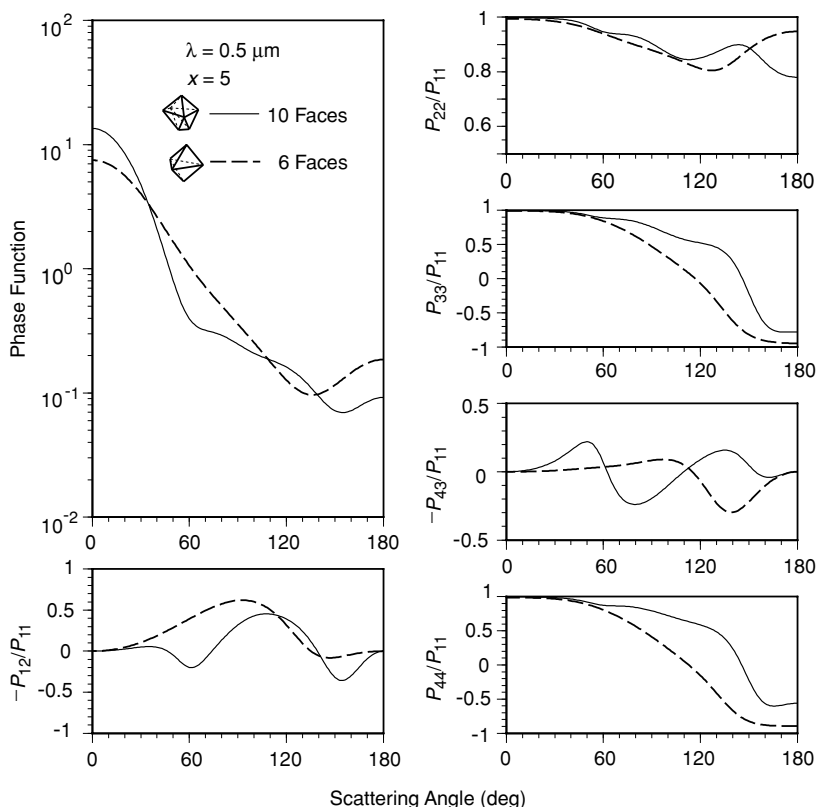
Figure 5.27   Scattering phase matrix elements for two types of randomly oriented dust aerosols. The size parameter used is 5 and the real and imaginary parts of the refractive index are 1.53 and 0.008, respectively, for a wavelength of 0.5 μm (data taken from Yang and Liou, 2000).

water-soluble, soot, oceanic, sulfate, mineral, water, and organic materials. To understand the scattering characteristics of aerosols, various representative aerosol geometries and inhomogeneous compositions are defined for light-scattering computations based on the FDTD method.

The left panels of Fig. 5.27 illustrate the phase function and the degree of linear polarization at $\lambda = 0.5\,\mu$m for two randomly oriented dustlike aerosol shapes, one with 10 faces and the other with six, constructed by a computer graphics program. The size parameters of these irregular aerosols are specified in terms of the dimensions of their peripheral spheres. Although the two polyhedrons have the same size parameters, the particle with 10 faces scatters more energy in the forward direction than its 6-face counterpart, because the volume of the former is larger. However, the reverse is true for backscattering. Note that the ratio of the extinction cross sections for these two aerosol shapes is 3.92. Dustlike aerosols are absorptive in the visible wavelength as indicated by the single-scattering albedos of 0.9656 and 0.9626 for the two polyhedral

geometries with 10 and 6 faces, respectively. The other panels of Fig. 5.27 show the remaining scattering phase matrix elements associated with the polarization state of the scattered wave. It is evident that the structure of aerosol geometry has a substantial impact on the polarization configuration.

Black carbon or soot aerosols generated from the incomplete combustion of fossil fuel and biomass burning can serve as condensation nuclei or become outside attachments to water droplets, a possibility relevant to cloud absorption. To illustrate the applicability of the FDTD method to light scattering by these types of aerosols, we present in Fig. 5.28 the phase functions and the degrees of linear polarization for four aerosol models defined in the diagram. The prime and double prime denote
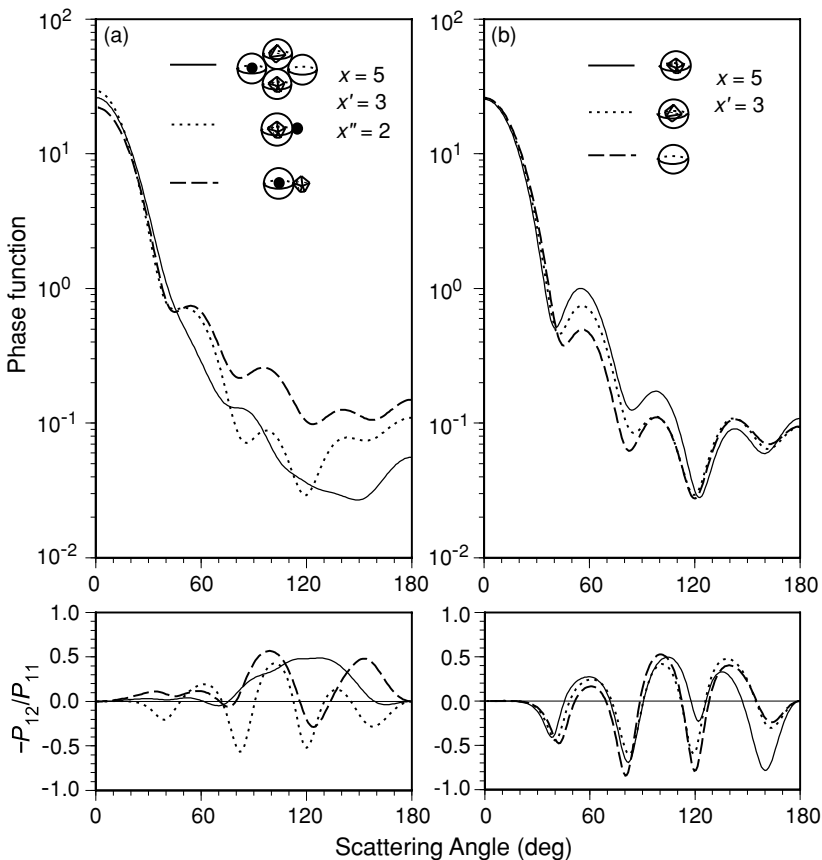

Figure 5.28 Phase functions and degrees of linear polarization for five aerosol geometric configurations using a wavelength of 0.5 μm. In (a), the size parameter for the largest water aerosol is 5, while the dust and spherical inclusions (or attachments) are 3 and 2, respectively. In (b), the inclusions are 10- and 6-face dust particles. The refractive indices used in the FDTD calculations are $(1.335, 1.0 \times 10^{-9})$ for water, $(1.53, 7.8 \times 10^{-3})$ for mineral dusts, and $(1.75, 0.45)$ for soot. Results for spheres computed from the Lorenz-Mie theory are shown in (b) for comparison purposes (data taken from Yang and Liou, 2000).

that the associated parameters are for mineral/dustlike and soot components, respectively, whereas the corresponding unprimed parameters are for water parts of the compounded particles. Polyhedral particles and sphere clusters produce smoother angular scattering patterns in comparison to the cases involving spheres with inclusions and/or attachments. In the case of the latter, the spherical parts of the compounded aerosols dominate the scattering properties. It is clear that the phase functions and polarization patterns of polyhedral and cluster aerosols are substantially different from those of homogeneous spheres, also shown in the diagram.

5.5.2   *T*-Matrix Method

The *T*-matrix method for light-scattering calculations is based on the expansion of the incident and scattered fields in vector spherical wave functions and relating these expansions by means of a *T*-matrix for computing electromagnetic scattering by single, homogeneous nonspherical particles. The *T*-matrix approach was initially introduced by Waterman (1971) and has been shown to be an efficient method for scattering calculations involving rotationally-symmetric nonspherical particles, such as spheroids, cylinders, two-sphere clusters, and Chebyshev particles (Mishchenko *et al.,* 2000). This method has been used for the analysis of light scattering by nonspherical aerosols, particularly with respect to the investigation of deviations of single-scattering results from those computed from the spherical assumption.

The essence of the *T*-matrix method begins with the expansion of the incident and scattered fields involving the scattering of a plane electromagnetic wave by a single particle in the forms

$$\mathbf{E}^i(\mathbf{R}) = \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \left[a_{mn} \mathbf{M}^*_{mn}(k\mathbf{R}) + b_{mn} \mathbf{N}^*_{mn}(k\mathbf{R}) \right], \tag{5.5.23a}$$

$$\mathbf{E}^s(\mathbf{R}) = \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \left[p_{mn} \mathbf{M}_{mn}(k\mathbf{R}) + q_{mn} \mathbf{N}_{mn}(k\mathbf{R}) \right], \tag{5.5.23b}$$

where $\mathbf{M}_{mn}$ and $\mathbf{N}_{mn}$ are certain vector spherical wave functions involving spherical Hankel functions; $\mathbf{M}^*_{mn}$ and $\mathbf{N}^*_{mn}$ are defined by similar functions, except spherical Hankel functions are replaced by spherical Bessel functions; a_{mn} and b_{mn} are expansion coefficients of the plane incident wave; and p_{mn} and q_{mn} are expansion coefficients of the scattered wave. Because of the linearity of Maxwell's equations and boundary conditions, the relationship between the scattered and incident coefficients must be linear and can be expressed by a *T*-matrix as follows:

$$p_{mn} = \sum_{\ell=1}^{\infty} \sum_{k=-\ell}^{\ell} \left[T^{11}_{mnk\ell} \, a_{k\ell} + T^{12}_{mnk\ell} \, b_{k\ell} \right], \tag{5.5.24a}$$

$$q_{mn} = \sum_{\ell=1}^{\infty} \sum_{k=-\ell}^{\ell} \left[T^{21}_{mnk\ell} \, a_{k\ell} + T^{22}_{mnk\ell} \, b_{k\ell} \right]. \tag{5.5.24b}$$

In matrix form, we write

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = \mathbf{T} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{T}^{11} & \mathbf{T}^{12} \\ \mathbf{T}^{21} & \mathbf{T}^{22} \end{bmatrix} \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}. \tag{5.5.25}$$

Equation (5.5.25) forms the foundation of the *T*-matrix approach. If the *T*-matrix for a given scatterer is known, the scattered field defined in Eq. (5.5.23b) can then be evaluated. Consequently, the amplitude matrix defined in Eq. (5.4.23a) can be obtained. The *T*-matrix depends only on the physical and geometric characteristics of the scattering particle, including the refractive index, size, shape, and orientation with respect to the incident light beam.

The general approach to determining the *T*-matrix is to find the field scattered by an object bounded by a closed surface in terms of an integral equation that involves the unknown surface field on the exterior of this surface. The determination of the internal field must then be defined by

$$\boldsymbol{E}^{\text{int}}(\mathbf{R}) = \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \left[c_{mn} \mathbf{M}_{mn}^*(mk\mathbf{R}) + d_{mn} \mathbf{N}_{mn}^*(mk\mathbf{R}) \right], \tag{5.5.26}$$

where m is the refractive index of the particle relative to that of the surrounding medium, and c_{mn} and d_{mn} are unknown coefficients. By using boundary conditions at the surface of the scatterer that require the continuity of the tangential components of the electric and magnetic fields, we obtain the following matrix equation:

$$\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}^{11} & \mathbf{Q}^{12} \\ \mathbf{Q}^{21} & \mathbf{Q}^{22} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}. \tag{5.5.27}$$

The elements of the $\mathbf{Q}$ matrix are surface integrals of products of the vector spherical wave functions that depend only on the particle's size, shape, and refractive index. Thus, by inverting this matrix equation, the unknown expansion coefficients of the internal field, $\mathbf{c}$ and $\mathbf{d}$, are expressed in terms of the known expansion coefficients of the incident field, $\mathbf{a}$ and $\mathbf{b}$. Moreover, from the boundary condition and the relationships among the incident, scattered, and surface fields, one can show that

$$\begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} = - \begin{bmatrix} \mathbf{Q}^{*11} & \mathbf{Q}^{*12} \\ \mathbf{Q}^{*21} & \mathbf{Q}^{*22} \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix}, \tag{5.5.28}$$

where the $\mathbf{Q}^*$ elements are also given by integrals over the particle's surface and depend only on the particle's characteristics. A comparison between Eqs. (5.5.25) and (5.5.28) leads to the solution of the *T*-matrix as follows:

$$\mathbf{T} = -\mathbf{Q}^* \mathbf{Q}^{-1}. \tag{5.5.29}$$

The preceding procedure for computing the *T*-matrix for single homogeneous scatterers in reference to the particle's coordinates is referred to as the *extended boundary condition method* (EBCM). The EBCM technique has been shown to have numerical stability problems in calculations for particles with very large real and/or imaginary parts of the refractive index, large size parameters (>20), and/or extreme geometries such as spheroids with large axial ratios. More efficient approaches that overcome the
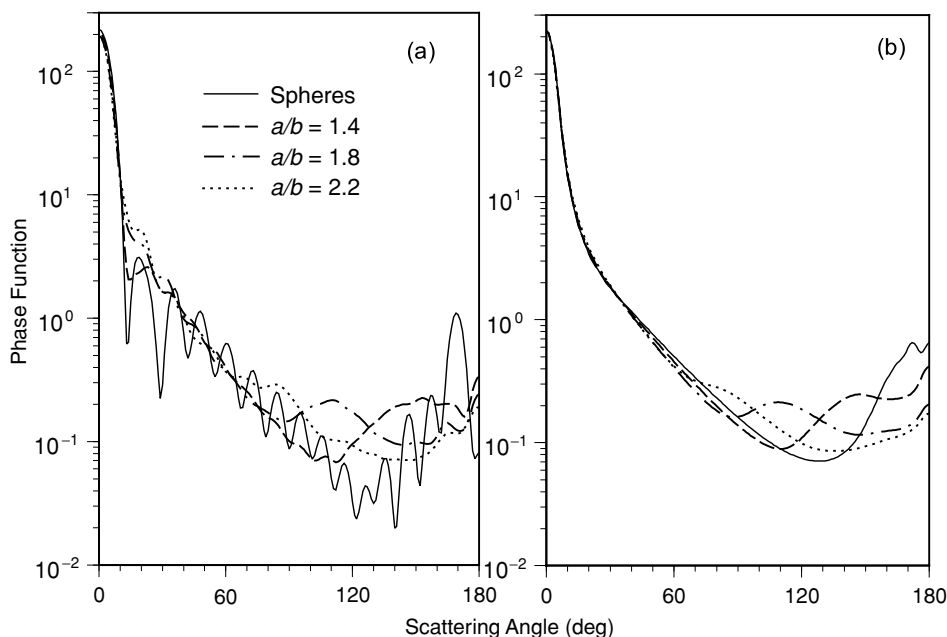
Figure 5.29   *T*-matrix computations of the phase function versus the scattering angle for monodisperse and polydisperse spheres and randomly oriented spheroids with a refractive index of $1.53 + 0.008i$. The wavelength is 0.443 μm. (a) Results for monodisperse spheres with a radius of 1.163 μm and surface-equivalent prolate spheroids with aspect ratios a/b, increasing from 1.4 to 2.2. (b) Similar computations but for a log-normal size distribution with a mean effective radius of 1.163 μm and an effective variance of 0.168 (data provided by Michael Mishchenko of the NASA Goddard Institute for Space Studies).

numerical instability problem in computing the *T*-matrix for highly elongated particles have also been developed. With computer coding improvements, the *T*-matrix method can be applied to symmetric homogeneous nonspherical particles with size parameters as large as 180. Interested readers may wish to refer to Mishchenko *et al.* (2000) for details.

Examples of *T*-matrix computations of the phase function as a function of the scattering angle for randomly oriented monodisperse and polydisperse spheroids and spheres are presented in Fig. 5.29. Panel (a) displays results for spheres with a radius of 1.163 μm and surface-equivalent prolate spheroids with aspect ratios ranging from 1.4 to 2.2. Panel (b) shows similar computations, but for a log-normal size distribution with an effective radius of 1.163 μm, illustrating the deviation in the phase function of nonsphericity from the spherical assumption.

Finally, comparisons of the phase functions of spheres computed from the Lorenz–Mie theory, of spheroids computed from the *T*-matrix approach, and of convex and concave particles computed from the FDTD method are presented in Fig. 5.30. These shapes are used to resemble aerosols with a refractive index of $1.38 + i3.9 \times 10^{-9}$ and a maximum size parameter of 10 with three aspect ratios. For $a/b = 1$
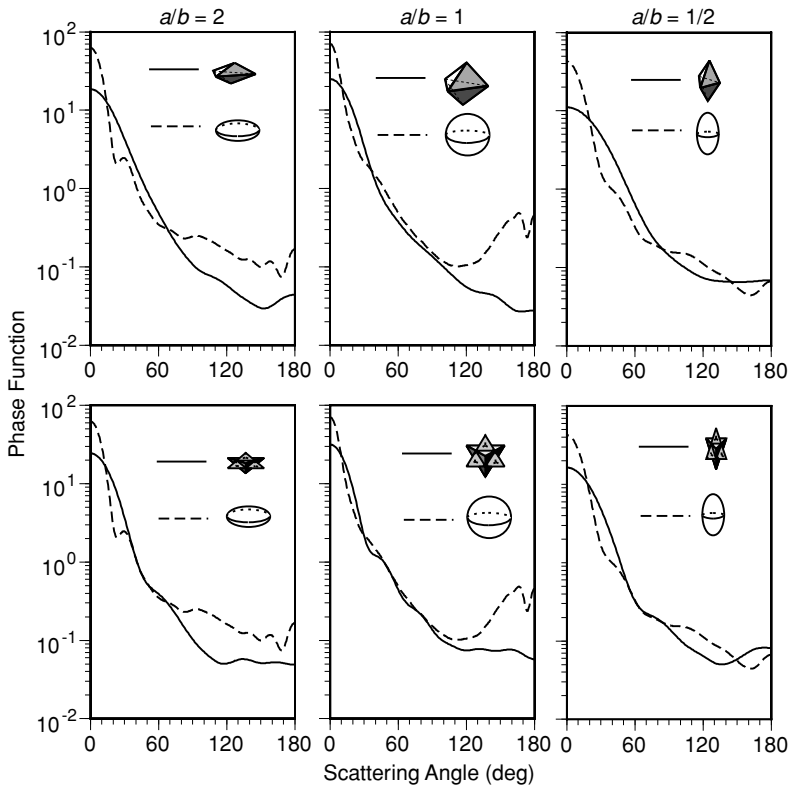
Figure 5.30   Comparison of the phase functions of oceanic aerosol particles with various shapes. For the spherical case, i.e., $a/b = 1$, the power law size distribution is employed to smooth out the resonant fluctuations. The maximum size parameter used is 10.

(a is the semimajor axis and b is the semiminor axis), the phase functions of nonspherical particles are substantially smaller than those of spheres in the scattering angular region $120°$–$180°$; in particular, the differences in the results for spherical and nonspherical particles are pronounced in the backscattering direction. For $a/b = 2$ and $a/b = 1/2$, we also see substantial differences between smooth spheroids and irregular convex and concave particles. Since the phase function is the critical parameter that is required in the satellite retrieval of aerosol optical depth (Section 7.3.3), selection of a proper and reliable value for each aerosol type must be undertaken to ensure the accuracy and precision of the retrieval.

5.5.3   Note on Light-Scattering Measurements for Nonspherical Aerosols

Development of integrating nephelometers for measurement of the scattering cross section of aerosols has been a subject of considerable research in the past (Heintzenberg and Charlson, 1996). The successful development of this instrument is significant

because if an accurate and reliable extinction cross section can be determined, the absorption cross section and hence the single-scattering albedo can be obtained to assess the climatic radiation forcing produced by aerosols.

Measurements of light scattering by aerosols as a function of wavelength and/or angular distribution for sizing purposes based on Lorenz–Mie results have been an interdisciplinary research subject, particularly related to instrument technology. However, angular scattering measurements of nonspherical aerosols, including polarization, have been extremely limited. The optical technique generally uses a laser beam and a sample of known aerosol types generated in the laboratory. Early laser scattering experiments with measurements of the scattering phase matrix have been reported by Holland and Gagne (1970) for randomly oriented microsized quartz crystals with flat platelike shapes. More recently, laboratory measurements of the scattering and polarization properties of aerosol particles have been presented by Volten *et al.* (1999). In angular scattering experiments, a He–Ne laser beam with a wavelength of 0.633 μm is usually used as a light source, which passes through a linear polarizer and travels through the electro-optic modulator, which can be rotated about its long axis. The light scattered in a particular direction passes through a quarter-wave plate and an analyzer and is then recorded by a photomultiplier detector. In this manner, the four-by-four scattering phase matrix can be measured. The sample aerosols are disturbed by air motion such that random orientation is anticipated. Thus, there are only six independent matrix elements. The scattering angular range covered is generally limited by the detector size with typical values ranging from 5° to 175° (see also Sections 5.4.4 and 5.3.3). The reliability of the experimental setup can be tested by employing spherical particles for which the measured data can be checked against results computed from the Lorenz–Mie theory.

An example of the scattering phase matrix elements for three types of mineral aerosols, feldspar, red clay, and quartz, with mean sizes 3.0, 5.1, and 9.7 μm, respectively, is shown in Fig. 5.31. Except for quartz, the refractive indices are basically unknown. The phase function (P_{11}) is normalized to 1 at the scattering angle of 30° in this presentation because of the lack of measured data from the scattering angle ranges 0–5° and 175–180°. The phase functions for these three aerosol types are extremely similar. The degree of linear polarization ($-P_{12}/P_{11}$) shows that the smallest particles have the largest maximum value. Close to the backscattering direction, negative polarization is seen. The remaining elements display different patterns for red clay than for the two other aerosol types, presumably because of its absorption property. A precise determination of the composition, size, and shape of aerosols based on scattering and polarization measurements, particularly from space, is critically important in order to reduce the uncertainties of the role aerosols play in global climate change.

As a result of the difficulty in manufacturing micrometer-sized aerosol particles with known optical properties and obtaining precise orientation information to compare to theoretical results, the microwave analog technique has been employed for light-scattering measurements involving particles with small size parameters. This technique usually involves incident microwave radiation from a transmitting antenna that passes a polarizer and is scattered by the object. The scattered radiation then
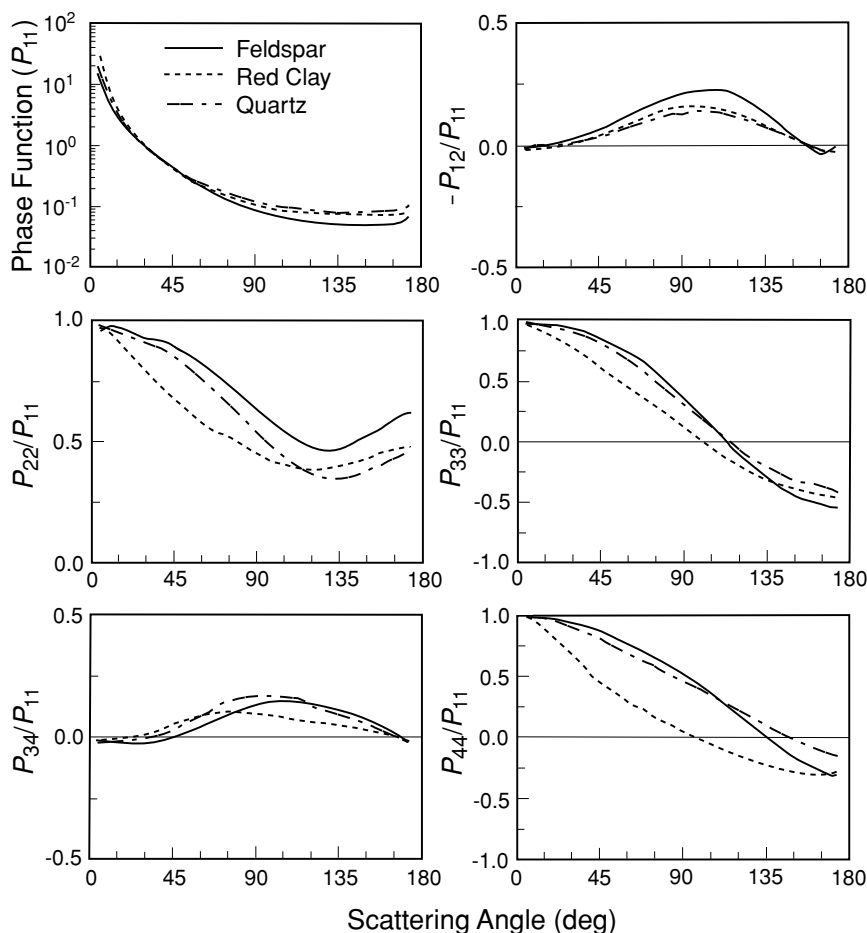
Figure 5.31 Measured scattering matrix elements versus scattering angle in a laboratory setting using a laser beam with $\lambda = 0.633\ \mu$m. The solid, dotted, and dash-dotted lines correspond to the results for feldspar, red clay, and quartz, respectively (data taken from Hovenier, 2000).

goes through the other polarizer and is measured by a receiving antenna. The method allows wide coverage of scattering angles, including the exact forward direction, as well as a greater degree of control of the object's size, shape, and orientation than conventional optical measurements. Figure 5.32 illustrates an example of phase function comparison between the theoretical results computed from the FDTD method and the experimental data derived from a microwave analog experiment. The measurements involved randomly oriented convex and concave particles with a refractive index of $m = 1.5 + i0.005$ and size parameters ranging from 5.9 to 17.8. In the FDTD calculation, a mixture of 50% convex and 50% concave particle shape was used. The phase function for this combination of particle shape appears to match the experimental results. Also shown for comparison is the phase function for the equivalent volume
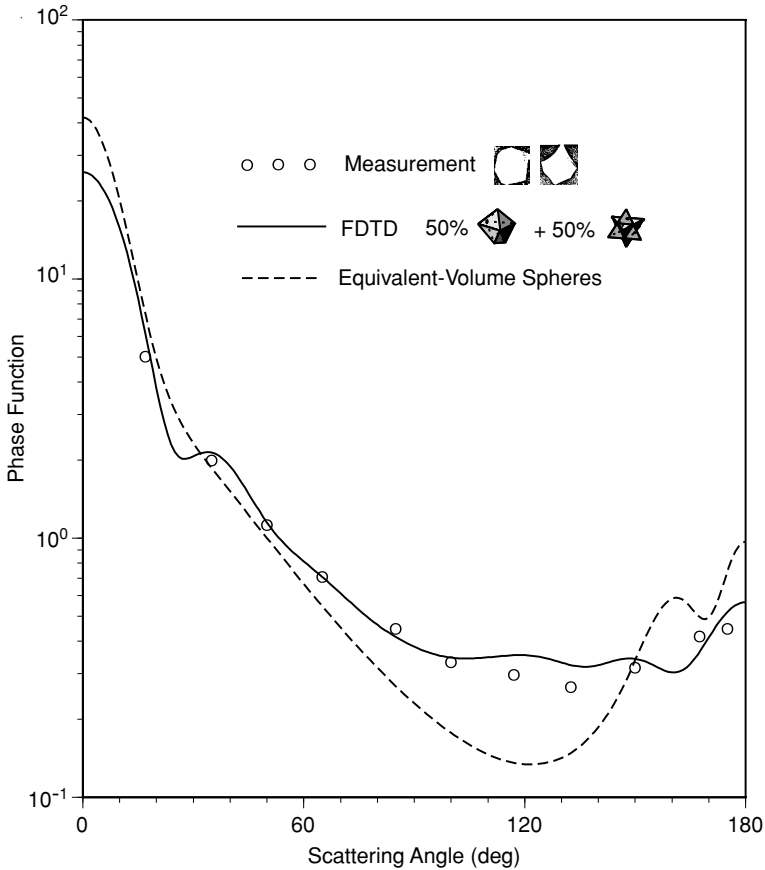
Figure 5.32　Phase function measured by the microwave analog technique (Zerull, 1976) and computed from the FDTD method for randomly oriented convex and concave particles with a refractive index $m = 1.5 + i0.005$ and a size parameter range from 5.9 to 17.8.

sphere computed from the Lorenz–Mie theory. Large deviations are evident between spherical and nonspherical particles, particularly in side scattering directions and the directions associated with the sun–satellite geometry.

Exercises

5.1　Based on the definitions of $\mathbf{M}_\psi$ and $\mathbf{N}_\psi$ in Eqs. (5.2.23) and (5.2.24), show that

$$\nabla \times \mathbf{N}_\psi = mk\mathbf{M}_\psi,$$

　　and prove that

$$\nabla^2\mathbf{N}_\psi + k^2m^2\mathbf{N}_\psi = 0, \quad \nabla^2\mathbf{M}_\psi + k^2m^2\mathbf{M}_\psi = 0.$$

5.2   From the radial component of the magnetic vector

$$H_r^i = e^{ik\cos\theta}\sin\theta\sin\phi = \frac{i}{k}\left[\frac{\partial^2(rv^i)}{\partial r^2} + k^2(rv^i)\right],$$

show that

$$rv^i = \frac{1}{k}\sum_{n=1}^{\infty}(-i)^n\frac{2n+1}{n(n+1)}\psi_n(kr)P_n^1(\cos\theta)\sin\phi.$$

5.3   The electric and magnetic field vectors in a homogeneous medium satisfy the following vector wave equation:

$$\nabla^2\mathbf{A} + k^2m^2\mathbf{A} = 0.$$

If ψ satisfies the scalar wave equation

$$\nabla^2\psi + k^2m^2\psi = 0,$$

(a) show that vectors $\mathbf{M}_\psi$ and $\mathbf{N}_\psi$ in cylindrical coordinates (r, ϕ, z) defined by

$$\mathbf{M}_\psi = \nabla \times (\mathbf{a}_z\psi), \quad mk\mathbf{N}_\psi = \nabla \times \mathbf{M}_\psi$$

satisfy the vector wave equation, where $\mathbf{a}_z$ is a unit vector in the z direction.
(b) Also prove that

$$\mathbf{E} = \mathbf{M}_v + i\mathbf{N}_u, \quad \mathbf{H} = m(-\mathbf{M}_u + i\mathbf{N}_v)$$

satisfy the Maxwell equations, where u and v are solutions of the scalar wave equation. Write out the expressions for $\mathbf{E}$ and $\mathbf{H}$ in terms of u and v.

5.4   The scalar wave equation in cylindrical coordinates is given by

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\psi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\psi}{\partial\phi^2} + \frac{\partial^2\psi}{\partial z^2} + m^2k^2\psi = 0.$$

Utilizing the method of separation of variables, show that the solution can be written as

$$\psi_n(r, \phi, z) = e^{i\omega t}Z_n(jr)e^{in\phi}e^{-ihz},$$

where h is an arbitrary constant, n is an integer, $j = (m^2k^2 - h^2)^{1/2}$, and Z_n is any Bessel function of order n.

5.5   Derive the explicit form of the transformation matrix $\mathbf{F}$ associated with the Stokes parameters in terms of $S_j(j = 1, 2, 3, 4)$, defined in Eq. (5.4.22).

5.6   Using the Fresnel formulas, show explicitly that the transmitted and reflected portions of the energy for the two polarization components are conserved. Also compute the incident angle θ_i at which $R_l = 0$. This angle is called the polarizing or *Brewster* angle, under which the electric vector of the reflected light has no component in the plane of incidence.

5.7   The refractive indices of water at the wavelengths of red (0.656 μm) and violet (0.405 μm) light are 1.332 and 1.344, respectively. (a) Describe the color

sequence of the corona. Find the radius of cloud particles that produce a sec-
ondary white corona with a 10° angular radius about the sun. (b) Describe the
mechanism of cloudbow formation. Find the scattering angles for the primary
and secondary cloudbows at these two wavelengths. (c) Find the angular radii
of the rings of the halos formed by prism angles of 60 and 90°. What will be the
widths of the rings?

5.8   When absorption is involved, the refractive index $m = m_r - im_i$. Prove that the
Fresnel reflection coefficients are given by

$$|R_r|^2 = \frac{(\cos \theta_i - u)^2 + v^2}{(\cos \theta_i + u)^2 + v^2},$$

$$|R_l|^2 = \frac{\left[(m_r^2 - m_i^2) \cos \theta_i - u\right]^2 + (2m_r m_i \cos \theta_i - v)^2}{\left[(m_r^2 - m_i^2) \cos \theta_i + u\right]^2 + (2m_r m_i \cos \theta_i + v)^2},$$

$$u^2 = \frac{1}{2}\left\{m_r^2 - m_i^2 - \sin^2 \theta_i + \left[(m_r^2 - m_i^2 - \sin^2 \theta_i)^2 + 4m_r^2 m_i^2\right]^{1/2}\right\},$$

$$v^2 = \frac{1}{2}\left\{-(m_r^2 - m_i^2 - \sin^2 \theta_i) + \left[(m_r^2 - m_i^2 - \sin^2 \theta_i)^2 + 4m_r^2 m_i^2\right]^{1/2}\right\}.$$

To derive these equations, let $m \cos \theta_t = u + iv$ in Eq. (5.3.23) and use the law
of refraction $\sin \theta_i = m \sin \theta_t$.

5.9   From the definition of the Stokes parameters in Eq. (5.2.103), show that the
scattering phase matrix for Rayleigh scattering [Eq. (3.3.6)] is given by

$$\mathbf{P} = \frac{3}{2}\begin{bmatrix} \frac{1}{2}(1 + \cos^2 \Theta) & -\frac{1}{2}\sin^2 \Theta & 0 & 0 \\ -\frac{1}{2}\sin^2 \Theta & \frac{1}{2}(1 + \cos^2 \Theta) & 0 & 0 \\ 0 & 0 & \cos \Theta & 0 \\ 0 & 0 & 0 & \cos \Theta \end{bmatrix}.$$

5.10 Consider a group of ice plates randomly oriented horizontally. (a) Show that
the incident angle of a geometric ray θ_i is related to the elevation angle ε_i
and azimuthal angle ϕ_i by $\cos \theta_i = \cos \varepsilon_i \cos \phi_i$. (b) Prove that the general Snell
law can be expressed by $m_r \sin \varepsilon_t = \sin \varepsilon_i$, and $m_r(\cos \varepsilon_t / \cos \varepsilon_i) \sin \phi_t = \sin \phi_i$,
where (ε_t, ϕ_t) are corresponding refracted angles and m_r is the real index of
refraction. (c) Find the angular distance between the sun and two bright spots on
its plane, referred to as *sundogs,* using a sun's elevation angle of 30°. Sketch a
diagram of sundogs showing the angular and azimuthal distances and the width
of red and violet colors.

5.11 The projection of a hexagonal column onto a horizontal plane is given by a
rectangle. Based on the Fraunhofer diffraction theory presented in Eq. (5.3.3),
show that the diffraction pattern for this aperture is proportional to

$$\frac{\sin^2(ak \sin \theta \cos \phi)}{(ak \sin \theta \cos \phi)^2} \cdot \frac{\sin^2[(L/2)k \sin \theta \sin \phi]}{[(L/2)k \sin \theta \sin \phi]^2},$$

where L is the length of the column, a is the half-width, and other notations are as defined in Eq. (5.3.3).

5.12 From the vector equation for geometric ray-tracing, show the individual components for external reflection, two refractions, and one internal reflection based on the directional cosine in polar coordinates.

5.13 Derive the adjusted real and imaginary refractive indices given in Eqs. (5.4.16 a,b) from Eqs. (5.4.9c), (5.4.13), and (5.4.14).

5.14 Show that the extinction cross section for a sample of randomly oriented hexagonal ice crystals of the same size is equal to half of the surface area of a hexagonal cylinder. In carrying out this exercise, first derive the following expression for the geometric cross section area of an arbitrarily oriented hexagon:

$$G(\alpha, \beta) = 3\sqrt{3}a^2 \sin\alpha/2 + 2aL \cos\alpha \cos(\pi/6 - \beta),$$

where α and β are orientation angles of a hexagon, L is the length, and $2a$ denotes the width.

5.15 Consider a sample of randomly oriented ice particles each of which has a plane of symmetry and to which the law of reciprocity is applicable. (a) Derive the six relationships for the scattering phase matrix elements by changing the signs of (S_3, S_4), in Eq. (5.4.22a). (b) Derive the additional four relationships by changing the signs of (U, V) in Eq. (5.4.29). (c) Prove that Eq. (5.4.32) is true.

Suggested Reading

Born, M., and Wolf, E. (1975). *Principles of Optics*. Pergamon Press, New York. Chapter 13, Section 13.5, gives a tractable approach to the solution of the vector wave equation for a homogeneous sphere utilizing the Debye potentials up to the point of the derivation of the scattered electric and magnetic fields. The book also presents an elegant treatment of basic and advanced material regarding electromagnetic scattering, geometric optics, and diffraction.

Greenler, R. (1980). *Rainbows, Halos, and Glories*. Cambridge University Press, Cambridge, U.K. A beautiful book that provides fascinating color pictures of various types of halos and arcs.

Kerker, M. (1969). *The Scattering of Light and Other Electromagnetic Radiation*. Academic Press, New York. Chapter 3 gives a comprehensive history of the solution for scattering by a homogeneous sphere. This book also provides a practical discussion of the scattering of light by spheres and cylinders in conjunction with the particle-size determination from scattering information.

Liou, K. N. (1992). *Radiation and Cloud Processes in the Atmosphere: Theory, Observation, and Modeling*. Oxford University Press, New York. Chapter 5 contains pertinent material and the theoretical foundations of light scattering by water droplets and ice crystals.

Mishchenko, M. I., Hovenier, J. W., and Travis, L. D., eds. (2000). *Light Scattering by Nonspherical Particles,* Academic Press, San Diego. This monograph provides a comprehensive and up-to-date overview of the subject of light scattering by nonspherical particles and its applications to science and technology.

Stratton, J. A. (1941). *Electromagnetic Theory.* McGraw-Hill, New York. Chapters 7 and 9 present advanced and organized materials for the solution of vector wave equations in spherical coordinates. (Stratton's text was the first to present this material.) Basic Bessel, Neumann, and Hankel functions and Legendre polynomials are also discussed in some detail in the book.

van de Hulst, H. C. (1957). *Light Scattering by Small Particles.* Wiley, New York. Chapters 8, 9, and 12 provide authoritative analyses of diffraction, Lorenz–Mie scattering theory, and geometric optics, respectively. In Chapters 2–4, fundamental scattering equations are derived by means of physical insight and postulations.