1

Introduction

Bright cumulus clouds on a summer afternoon contrast vividly with a darker blue sky. Later in the day, rainbows may add multicolored arches during a rainstorm. Other colors in nature are the dark green of forest foliage and the red and orange hues of the Grand Canyon. On high mountains or in deserts, dark patches in the bright band of the Milky Way can be seen on clear nights. Chimney smut blackens all it touches, and iridescent opal shimmers with colors. All these observations of light and color and many more result from absorption and scattering of light by small particles. We use "light" loosely for electromagnetic radiation in and near the visible spectrum but shrink from extending it to radio wavelengths. For most of recorded history, light was a *perception*, not known to be stimulated by radiation in a narrow wavelength interval. Genesis 1:3 does *not* read, God said, "Let there be electromagnetic radiation between 400 and 700 nm" (despite what we have seen on T-shirts). Light acquired the secondary meaning of this radiation itself. Sound underwent a similar transformation from a perception to its physical stimulus.

Because light scattering theories, measurements, and applications are too broad to be treated in one book, we limit ourselves mostly to *single* scattering by particles described by macroscopic electromagnetic theory. Single scattering means that particles in a collection are excited to scatter mostly by an external source and negligibly by each other, unlike for some of the preceding examples. To understand the brightness of clouds requires theories of *incoherent* multiple scattering outside the scope of this book. The qualifier "incoherent" signals that phases of waves scattered by cloud droplets *may* be ignored because their positions are uncorrelated, and the inverse cube root of their *concentration* (see Zeng et al., 2014) is more than about 30 times the lateral coherence length of sunlight (Sec. 4.4.8). Incoherent scattering is sometimes called *independent scattering*, in contrast with *dependent* or *cooperative* scattering. But cloud droplets illuminated by sunlight are *not* independent because the light scattered by each *depends* on the light scattered by its neighbors unless the cloud is very thin. Although *independent* was used by, and sometimes attributed to, van de Hulst (1957, Sec. 1.21), he subsequently replaced it with *incoherent* (1980, p. 20).

A *single* cloud droplet is a *coherent* scatterer. Because the positions of its molecules are correlated, scattering *must* be treated with theories in which phases are not ignored, although we need not explicitly consider scattering by individual water molecules. But ordinary clouds composed of many droplets are incoherent scatterers, an example of *classical* decoherence (for *quantum* decoherence, see Schlosshauer, 2019). Closely packed silica spheres (opal) with diameters of the order of 0.1 μm (Sanders, 1964; Sanders and Murray, 1978) are cloudy but display iridescent colors (whence the term opalescent); their optical properties at visible wavelengths lie between those of liquid water (homogeneous on the wavelength scale) and atmospheric clouds (inhomogeneous on the wavelength scale). The colors of rocks and foliage originate from the interaction of sunlight with many densely packed centers of absorption and scattering. Detailed treatments of these phenomena begin with the simpler ones we treat. To avoid writing molecules and atoms, in the rest of this book, molecules include atoms unless stated otherwise.

To determine coherent scattering by particles, we must know their positions because they determine phase differences between scattered waves. Reflection and transmission by a homogeneous slab (Sec. 2.7) are problems for which we know the positions of all small volume elements excited to scatter by an incident wave. The only position variable is distance into the slab. All its separate parts move in concert if pushed slightly because they are connected, and hence stay *correlated*, unlike droplets in a cloud. We can walk on an ice floe but not a cloud—except metaphorically.

Theories of coherent and incoherent multiple scattering are sometimes distinguished by calling the former multiple scattering of *waves* (or fields) and the latter multiple scattering of light corpuscles (*photons*). *Radiative transfer* or *transport* usually denotes the domain of multiple-scattering theories in which the electromagnetic field equations are bypassed and phase differences are ignored, because of which we need not know the positions of scatterers, only their concentration. Such theories are sometimes said to be valid if the wavelength is small compared with the average nearest-neighbor distance between randomly separated scatterers. Although this condition is *sufficient* and is satisfied by ordinary clouds from visible to far infrared wavelengths, it is not *necessary*. Wavelengths of solar radiation are ≈ 100 times *larger* than the average separation between air molecules in the lower troposphere, and more than twice the mean free path between intermolecular collisions. Despite this, radiative transfer theory agrees with measurements of solar radiation scattered by Earth's molecular atmosphere and laboratory measurements of scattering by gases at ordinary temperatures and pressures (Sec. 5.8). Transmission of sunlight by Earth's molecular atmosphere is reduced much less by coherent scattering (specular reflection) than by incoherent scattering because the refractive index of air is ≈ 1, and the average (incoherent) scattering mean free path is of the order of kilometers (Sec. 11.5). This is the reverse of transmission by solid or liquid films with refractive indices appreciably different from 1 but much thinner than the incoherent scattering mean free paths (Sec. 5.8). For the meanings of mean free path, see Sections 3.4.1 and 5.8. For more about the limits of radiative transfer theory, see Section 2.4.

Monte Carlo methods applied to radiative transfer in incoherently scattering media depict light as streams of localized massless corpuscles endowed with energy and momentum, which are scattered and absorbed according to statistical prescriptions.

The average nearest-neighbor distance between particles, relative to their size, does not provide a sufficient condition, *independent of wavelength*, for incoherent multiple scattering. Sugar grains in a bowl are touching, but at visible wavelengths sugar is an incoherent multiple-scattering medium similar to a cloud, whereas at centimeter wavelengths it is a coherent multiple-scattering medium ($\approx 10^3$ grains cm^{-3}) similar to a chunk of rock candy.

Although radiative transfer is a theory of *real scalar* fields, it incorporates parameters such as scattering and absorption cross sections obtained either from measurements or theories of *complex vector* fields. These parameters are foreign imports from the wider electromagnetic world, not the domestic products of radiative transfer theory.

Scattering (single or multiple) by *random media* may denote scattering by discrete particles randomly distributed in space and time (e.g., clouds) or scattering by media with assumed continuous properties randomly varying in space or time or both (e.g., turbulent air).

Shifrin (1951, p. 12) notes that four length scales relative to the wavelength determine scattering by a collection of particles: molecular, particle, inter-particle, and overall size. We are concerned mostly with the second. Macroscopic electromagnetic theory eliminates the first by averaging; radiative transfer theory eliminates the third by further averaging.

1.1 What Is a Small Particle?

Particle is burdened with several meanings. An *elementary* particle (e.g., electron) has no structure but does have an invariant mass and charge. But Dehmelt (1990) argues that "the electron may have *size and structure*." The best current measured (ACME, 2018) upper limit of the *electric* dipole moment of the electron is $|p_e| < 1.1 \times 10^{-29}$ e cm, evidence that its charge e is confined within an extraordinarily small nearly spherical volume ($\approx 10^{-66}$ nm^3). For arguments that "a point-like particle with finite charge and…mass is a strict impossibility in classical [an important qualifier] physics," see Steane (2015a) and references cited therein.

A *composite subatomic* particle (e.g., proton) is composed of elementary particles (e.g., quarks) and is itself a component of atoms. A molecule is a bound aggregation of a fixed number of atoms (e.g., H_2O). All these entities are often called particles. Another meaning is a vanishingly small part of an idealized continuous medium, for example, the fictitious particles in fluid mechanics.

For our purposes, a particle is a *bound* aggregation of an indefinite number of molecules sufficient in quantity to have bulk *properties* such as permittivity and the *qualities* of solidity or fluidity. How many are required for an aggregation to qualify as a particle cannot be specified precisely. A single molecule does not possess these properties or qualities, which are *averages* over many molecules. Boundaries separating single molecules from molecular clusters, and molecular clusters from particles, are fuzzy. Consider N_m molecules per unit volume in a compact aggregation with area A and volume V. Only those molecules well within the particle are *not* in an atypical environment. The number of molecules near the surface is $N_m\ A\ d$, where d is a few molecular diameters. Thus, the fraction of molecules in an atypical environment is d/D, where $D = V/A$ is a characteristic linear dimension. A rough criterion for the maximum size of a compact aggregation to be considered a particle is $d/D \ll 1$. If we take d to be 1 nm, D must be >10 nm for an aggregation to have bulk properties. A possible exception is metallic particles (Sec. 12.1.7). Although macromolecules can be as large as colloidal particles, molecules are distinguished from particles by their individuality: a molecule deprived of one or more atoms changes its chemical identity. Nor do molecules grow, whereas particles (e.g., water droplets) can grow by accretion or shrink by net evaporation. Compact means that the environment of most of the atoms or molecules in the aggregation is not greatly different from what it would be in the bulk material. Thus, N_m is *not* independent of the composition of the particle nor even the wavelength of the illumination. Molecules on or near its surface are always in an environment unlike that within the particle. This is shown nicely by Draine (1988, Fig. 2), who addressed the question of how many finite dipoles are necessary to adequately approximate a sphere continuous on the molecular scale and how this depends on its refractive index.

By an unqualified particle, we usually mean matter within a bounded volume, although we consider infinitely long cylinders (Sec. 8.3) and slabs of finite thickness but infinite in lateral extent (Sec. 2.7) to be particles. And we sometimes compare absorption and scattering by such idealized particles with that by particles bounded in three dimensions.

The qualifier "small" is ambiguous. Small and large have no meaning unless a scale is specified or implied. Although in optics the scale often is the wavelength, "small" to us does *not*, in general, mean small compared with the wavelength. Most of the particles we consider lie in the nanometer- to centimeter-size range and are illuminated by radiation from ultraviolet to radio wavelengths. The "small" particles we treat are *not* elementary (subatomic) particles or (usually) individual atoms or molecules.

An aggregation of particles, not necessarily touching, in *fixed* relative positions can itself be considered a particle if the illumination is ideally coherent (e.g., a plane harmonic wave). Real sources place limits on the overall size of an aggregation for it to be considered a particle. We would not consider an ordinary cloud illuminated by *sunlight*, even with all its droplets frozen in place, to be a single particle because there is no fixed phase relationship between the waves scattered by the individual droplets. The finite *lateral* (or *transverse*) *coherence length* (Secs. 1.3, 4.4.8, and 7.2.1) of the illumination gives us a *sufficient* condition for whether a particle is small enough that scattering by it is described well by theories in which the source is ideally coherent. By this criterion, a billiard ball ($\approx$6 cm) illuminated by 1-cm radiation would be reckoned as "small" even though it would not feel small if dropped on a bare foot.

This classification is not as absurd as it might seem, because the ball scatters like fog droplets illuminated by sunlight, both described by the same theory because they are smaller than the lateral coherence length of the illumination.

Because in *theories* of scattering the illumination is often explicitly (or implicitly) a structureless plane wave unbounded in lateral extent, it is easy to forget that it *never* is in *experiments*. A laser beam is highly coherent and localized laterally, and hence has lateral structure.

Green and Lane (1964, p. 3) note that attempts to classify particles according to their nature, origin, and size have not been completely successful. Liquid particles usually are called *droplets* or *drops* depending on size; typical raindrops are about 100 times larger than typical cloud droplets. *Fog* often means a cloud near the ground, with droplets in the smallest size range and sufficiently high in concentration to markedly diminish visual range. Haze usually means an aerial suspension of solution droplets limited by the sub-saturated humidity of their environment from growing to cloud droplet size. *Dust* or *powder* often means solid, irregularly shaped particles originating from the "mechanical disintegration of matter" and suspended in a gas, but interstellar dust often denotes the much smaller particles (or *grains*) between stars. *Smokes* may be particles, usually solid, formed by combustion or other reactions in which molecules combine to form aggregations. The term *aerosol* was coined to mean particles suspended in a gas and sufficiently small to be stable with respect to sedimentation but it has been extended to mean any suspension of particles in air (Green and Lane, 1964, p. 5). Without qualification, a *sol* is a suspension of particles in a liquid, and it is called a *hydrosol* if the liquid is water. Both aerosols and hydrosols include the suspending medium. By *colloidal* particles is meant those ranging from nearly molecular size to the smallest size "that can be resolved by the ordinary microscope," between about 1 and 500 nm (Hauser, 1939, p. 25). A *bubble* is usually a void in an otherwise homogeneous liquid or solid. But *any* particle with a refractive index less than that of the surrounding medium is optically a bubble.

Because of the lack of terms universally agreed upon, we use the nonspecific term particle as much as possible and leave it to metaphysicians to decide if particles belong to a fog, dust cloud, or smoke. Although these and other terms may have definite meanings within a field, they are likely to change across boundaries. The term *obstacle* may denote a molecule or a particle but connotes an object large enough to trip over.

By clouds without qualification, we mean ordinary clouds of water droplets or ice particles in Earth's atmosphere or inaccessible clouds in planetary atmospheres (e.g., sulfuric acid clouds on Venus), to be distinguished from dust or smoke clouds, particle suspensions in air that in the aggregate have ill-defined shapes and sizes. Clouds are ephemeral and, appropriately, nebulous.

Whatever its composition and origin, a *nanoparticle* (or *ultrafine* particle) has a diameter in the range of 1–100 nm.

1.2 Scattering, Emission, and Absorption as Observable Phenomena

Light scattering was observed long before the nature of light and matter was understood. An empirical view of scattering, neither enlightened nor encumbered by theory, is as follows. Consider a point P in the path of a beam of light. A detector (e.g., a human eye) is outside the beam at an arbitrary location and directed at P. With the beam turned off, a particle is placed at P. The detector registers zero. Turn on the beam and quickly turn it off. The detector responds to a flash of light from the particle. It is reasonable to say that an *incident* pulse of light is a *cause* that precedes in time an *effect*: the detector's response to the flash of *scattered* light. But it is futile to ask if this flash is a fragment of the incident light, because we cannot answer this by experiment. We can observe only that the incident light excites the particle to produce a flash (i.e., scattered light) at the detector.

If we turn off the beam and place a particle at P, a suitable detector may respond to light *emitted* by the particle, *not* excited by an external source. With one exception (Sec. 4.7) we ignore emission. The distinction between scattering excited by an external source and spontaneous emission is sometimes muddied, and scattered light sometimes is denoted as emitted light. But the distinction is real and should be made. Because light pulses don't carry identity cards, the notion of "re-emission" is meaningless.

Radiation and scattering appearing in tandem implies a distinction. Scattering requires excitation by radiation, whereas (unqualified) radiation does not, as, for example, thermal emission or radiation from antennas driven by electric currents. But sometimes the distinction is blurred. The electromagnetic energy within the confines of a particle illuminated by a narrow pulse rapidly increases, and the particle scatters while *directly* excited by the pulse. After it has moved on and the exciting field at the particle has dropped to zero, it continues to radiate for a time longer than the pulse duration, decreasing in amplitude with a decay time depending on particle characteristics (see, e.g., Garner et al., 2017a, Figs. 3, 5, 7). Even absent absorption, the particle's electromagnetic energy decreases, analogous to *reverberation*, the gradual fading away of sound after its source has become undetectable.

Consider a suspension of many identical particles. Turn on the beam for a brief moment. Flashes from each particle occur at slightly different times; their duration and magnitude are not the same as for an isolated particle because each particle is illuminated by a primary source and many secondary sources (light scattered by its neighbors). This is multiple scattering in the *time domain*. We cannot readily observe this temporal sequence because d/c, where d is the average distance between particles and c the free-space speed of light, usually is exceedingly small (picoseconds).

Then turn on the beam and leave it on. Because d/c usually is much smaller than the response time of detectors, the suspension *appears* to be continuously luminous because of an irresolvable sequence of events rapidly playing out again and again. No light sources are constant in time, but oscillate with periods of the order of femtoseconds (10^{-15} s) at visible and near-visible frequencies. If periodically time-varying sources are averaged over many periods, and if detector response times are much greater than these periods and d/c, descriptions of scattering need not consider time explicitly. This is why many scattering problems can be formulated and solved in the *frequency domain* (Sec. 2.3.1). But we become so accustomed to this fortunate state of affairs that we can easily forget time always lurking in the background. Moreover, we tend to think in the time domain—first A happens and then B—but solve problems in a domain with neither past nor future, only an eternal present. Yet illumination of a particle always has a beginning, following which the fields within and scattered by it become steady in a time interval so short as to be (almost) unobservable.

Again, consider a single particle illuminated by a beam. The detector responds to scattered light, and the *state* of the particle may change, manifested by an increase in its temperature or a phase change (e.g., solid to liquid) or a chemical reaction. This change of state, lumped under the general category of *absorption*, is all that can be observed. No identifiable thing or substance has been absorbed. Saying that light is absorbed by a particle is shorthand for saying that the interaction of the beam with the particle causes its state to change, often below the limits of detection.

Absorption of electromagnetic radiation, by analogy with the directly experienced absorption of water by a sponge, is figurative rather than literal because water is incorporated into the sponge's structure whereas absorbed electromagnetic radiation is *not* incorporated into the structure of matter.

A beam may be obtained from a light source plugged into a socket from which extend wires to a generator driven by an exercise wheel in a rat's cage. The rat eats sunflower seeds, converting their (chemical) potential energy into translational kinetic energy (the rat race), which in turn is converted into rotational kinetic energy of the wheel, which is converted into electrical power, which is converted into a beam that illuminates a particle, the temperature of which increases. What was absorbed? Sunflower seed potential energy, itself transformed solar energy originating from thermonuclear reactions? Translational kinetic energy of a panting rat? Rotational kinetic energy? Electrical energy? Radiant energy? A chain of several mediators results in a change of state of the particle. The beam is the *last* link in this chain, the *proximate cause* of the particle's change of state.

1.3 Detecting and Imaging

"To see" could mean to detect or to resolve or to image. The essence of imaging is a one-to-one mapping (never perfect) from object points to image points. A lens effects such a mapping. Although to be imaged particles must scatter or emit light, we restrict ourselves, with one exception (Sec. 13.1.4), to detection of scattered light in the following sense. If the photosensitive surface of a detector lies in the focal plane of a lens that collects light scattered by a particle, it is imaged on this surface. But the detected characteristic of the image is its *radiometric* magnitude, not its size, shape, and orientation. Nevertheless, an image of the particle is the scattering diagram of the lens (in its focal plane).

Although most images are formed with scattered light, not all scattered light results in resolvable images. The distinction is exemplified by the ultramicroscope developed a century ago by Zsigmondy and Siedentopf to detect colloidal gold particles. Today, ultramicroscope often connotes high magnifying power, but Zsigmondy's ultramicroscope was low power, and to him "ultra" meant below the limit of resolution. He could detect gold particles as small as 10 nm but not resolve them. Why? The superficial answer, they were "too small," won't do. On a clear night we can detect stars but cannot resolve them, and yet they are huge. Zsigmondy observed on a small *linear* scale what we observe on a vastly larger linear scale if we gaze at a star-filled sky because the *angular* scales are comparable. The angular width of a 10-nm particle at 5 cm is 2×10^{-7} rad, close to the angular width of the star Betelgeuse (Wolf, 2007, p. 8). The *Rayleigh criterion* (Rayleigh, 1879–80, §1) for the minimum angular separation between two resolvable points is $\theta \approx 1.22\lambda/D$, where λ is the wavelength and D is the diameter of a *perfect* (no aberrations or irregularities) lens focusing light onto a perfect detector. The angular resolution of the Hubble Space Telescope (mirror diameter 2.5 m) not limited by the atmospheric turbulence and turbidity that spoils the resolution of Earth-bound telescopes is about 2.4×10^{-7} rad (≈ 50 milliarcseconds), close to the theoretical limit but not good enough to resolve *details* of Betelgeuse.

Rayleigh's resolution criterion is the most well known but not the ultimate. den Dekker and van den Bos (1997) aver in a survey of resolution that because it "is not unambiguously defined, it is interpreted in many ways." They define "classical resolution limits" as "those exactly calculated by a mathematical model," noting that "Such images do not occur in practice" and that "Since Rayleigh's days, technical progress has provided us with more and more refined sensors." The "classical criteria produce resolution limits that are independent of any condition other than the size and shape of the imaging aperture and the wavelength of the light." The image of a point is not a point but rather a fuzzy *Airy disk* (Sec. 4.4.3).

Despite its familiarity in textbooks, Rayleigh's two-point criterion is a rule of thumb, "never claimed… [to be] a hard limit" (de Villiers and Pike, 2016).

Eyeglass lenses correct defects of vision; microscope lenses resolve fine details; telescope lenses make distant objects appear closer. Resolution unqualified usually means that of the (possibly augmented) human visual system. But it is possible to process photographic images to detect rainbows of third and higher orders (Sec. 7.2) that cannot be observed otherwise.

Resolution in microscopy is determined by the wavelength and the numerical aperture of the lens. Regardless of how resolution is defined or specified, it must include wavelength and optical system geometrical parameters. Resolution, magnification, and contrast are different but related concepts (see, e.g., Masters, 2020, Ch.1).

Zsigmondy illuminated gold particles with focused sunlight and observed them against a black background. Because of sufficient *contrast* with the suspending medium he could detect but not resolve them. Small metallic particles in water scatter light much more than insulating particles of the same size as solvent molecules or those in solution. Similarly, we can detect stars at night because of sufficient contrast but cannot resolve their shapes or features such as sunspots (stellar "freckles").

Before moving to prosaic matters, we quote Zsigmondy's (1909, pp. 134–135) poetic description of what he observed: "A swarm of dancing gnats in a sunbeam will give one an idea of the motion of the gold particles in the hydrosol of gold! They hop, dance, jump, dash together, and fly away from each other, so that it is difficult in the whirl to get one's bearing." Oscar Wilde quipped, "We are all in the gutter but some of us are looking at the stars." Zsigmondy was in his laboratory looking at stars—through his ultramicroscope.

In a typical experiment for detecting visible and near-visible light scattered by a particle, a lens, possibly limited by stops, is followed by a detector at a distance r from a particle of diameter d and illuminated by a primary source (e.g., laser beam) with wavelength λ. The particle is a secondary source; its *far field* (Secs. 4.4.3 and 7.2.4) is at distances *r at least* greater than $2d^2/\lambda$. A lens at $r > 5$ cm is in the far field of particles with $d < 100$ μm.

The pupil diameter of a human eye is 2–8 mm and depends on many factors (Watson and Yellott, 2012); "a 3-mm pupil usually produces optimum image quality" (Boynton, 1979, p. 79). For 5 mm, the minimum angular resolution according to the Rayleigh criterion is $\approx 10^{-4}$ rad. Actual is larger because of defects in the eye: a value frequently cited is one arcminute ($\approx 3 \times 10^{-4}$ rad). A particle of diameter 100 μm is barely resolvable at 25 cm, the *least distance of distinct vision* or *near point*, the average of a characteristic that varies among humans and increases markedly as they age (presbyopia). With a conventional optical microscope, the best we can do is resolve two points around one-third the wavelength apart, ten times larger than Zsigmondy's gold particles. For an unconventional lens that exploits multiple scattering to achieve 100-nm resolution, see van Putten et al. (2011).

A strictly time-harmonic field, unattainable in practice but the basis of much theory, is *completely* coherent in that the modulus of its normalized mutual coherence function as well as that of its normalized autocorrelation function is unity (Read, 1980, Ch. 5). Quasi-monochromatic fields (Sec. 2.10.1) are only *partially* coherent. Even if filtered by a narrow band-pass filter, sunlight or light from an incandescent lamp constrained by a small aperture (e.g., a pinhole) is not completely coherent but has a finite lateral coherence length. For a wavelength λ and a solid angle Ω subtended by the source at r, the lateral coherence length is $\lambda/\sqrt{\Omega}$; for a simple derivation for an extended source each point of which is uncorrelated with every other point, see Bohren and Clothiaux (2006, Sec. 3.4.2), also Section 4.4.8. For a rigorous derivation of essentially the same expression for sunlight see Agarwal et al. (2004).

The coherence length of an idealized point source is infinite. The smallest angular diameters of resolvable stars are of order one milliarcsecond (mas). For example, the angular diameter of Alpha Centauri B (Kervella et al., 2003) is about 5.9 mas, which corresponds to a coherence length at visible wavelengths of about 20 m.

A large water drop is both a particle and a lens (Sec. 7.2.4). Although not usually thought of as a particle, a lens is a nonspherical, negligibly absorbing particle much larger than the wavelengths of the illumination. The near-forward scattering pattern *of the lens* limits its resolution. But it scatters light in every direction. That in the backward hemisphere is called reflected light and usually ignored as a negligible loss.

If we rotate a lens-detector system on an arc at a constant distance from an illuminated particle we can measure its *scattering diagram*, the detector response for a range of directions, a characteristic of the particle but not its image. Although by analyzing a scattering diagram we *might* be able to infer the size and shape of a particle, we can determine them at a glance *if* it can be imaged. We discuss an example of imaging—glare spots—in Section 13.1.4.

1.4 Elastic, Quasielastic, and Inelastic Scattering

We usually assume *elastic* scattering: the frequency of the scattered radiation is that of the incident radiation. Elastic scattering is sometimes called *coherent* scattering. Scattering of radiation by an electron according to Thomson's theory (Sec. 9.1.2) is sometimes said to be coherent (also *unmodified*) because it predicts no frequency change, whereas scattering of radiation by an electron according to Compton's theory is sometimes said to be incoherent (also *modified*) because it predicts a small frequency change except in the forward direction. But incoherent may mean that the phases of waves scattered by individual atomic electrons are assumed to be random, and hence their total scattered *power* is additive.

Elastic scattering is more physically descriptive than coherent scattering and less likely to lead to confusion because it does not clash with the well-established definition of coherence as correlation between two or more sources of radiation (Forrester, 1956). In this sense scattering in a given direction by two or more scatterers may be elastic but not necessarily coherent. Coherence can be qualified by specifying the degree of correlation. Sometimes coherent has the more restrictive meaning of *zero* phase difference. We restrict coherent scattering to a definite phase relation between two or more scatterers, possibly parts of the same scatterer, a definition that implicitly includes the properties of the source. Coherence means the sticking together of entities, usually of similar nature. To add to the burden "coherent" must bear are coherent *arrays*, all elements of which are fixed relative to each other, coherent *sources* (the coherence of which is a matter of degree), coherent *detection*, and coherent *theories*. Even more reason to prefer elastic scattering to coherent scattering.

Defining coherence in terms of phase correlations is unexceptionable, but we cannot observe them directly, at least not easily. A definition in terms of *spatial* interference patterns is by Zernicke (1938): "*Two vibrations of light shall be called incoherent, if their superposition gives no visible interference*." Any departure from this incoherence is evidence of *partial* coherence to varying degrees. He defines the degree of coherence in terms of the visibility (essentially the contrast) of interference *fringes*. Because of his interest in microscopy (he invented the phase-contrast microscope), "visible interference" to him meant visible to the human eye. We can qualify his criterion by specifying the resolution of the detector. Interference may also refer to *temporal* patterns, the occurrence of beats when two beams of slightly different frequencies are superposed. *Any* two beams can interfere over a sufficiently short time or over a sufficiently small area, which may or may not be resolvable depending on the detector (see the end of Sec. 2.4).

Localized *dark* fringes capture our attention. Adding light to light to yield *brighter* fringes is hardly a cause for wonder. If you can't read this book by the light of one lamp, turn on another.

Superposition and interference are not synonymous, although the latter requires the former. Two or more electromagnetic fields may superpose (Sec. 2.1) but not interfere in that the time-averaged Poynting vector (Sec. 2.4) of the superposed fields is the sum of their individual time-averaged Poynting vectors independent of all the others. Any departure from this may be called interference, the usual meaning of this term in optics. For more about interference, see the end of Section 2.4 and Kipnis (1991, Sec. 1.4). See also the discussion of the Arago-Fresnel laws in Section 2.10.1.

Spatial interference fringes are not the only result of coherence. For example, bands in *spectral* transmission by thin films (Fig. 2.8) are an interference pattern resulting from the superposition of many coherent sources; every part of such films is excited by the incident wave and correlated with every other part. In a sense, a thin film is a beam splitter that splits an incident beam into a great many beams, which can interfere if the film thickness is less than the lateral coherence length of the source. For Planckian ("blackbody") radiation, this is of the order of the bandwidth (Mehta, 1963).

Coherent and monochromatic are *not* synonymous. The Sun, or any incandescent source such as a light bulb, is an incoherent source of polychromatic radiation with a small lateral coherence length but large enough to result in (some) interference patterns. Lasers are coherent sources of nearly monochromatic radiation.

Strict elastic scattering does not exist because matter is in motion on some scale, resulting in small but sometimes measurable Doppler shifts. *Quasielastic scattering* (also called *dynamic light scattering*) refers to small departures from strict elastic scattering (also called *static light scattering*), which can be measured and may be worth the effort to do so. Although a source may be nearly monochromatic with carrier frequency ω_c, the radiation scattered by scatterers moving relative to each other varies in time, and hence its Fourier transform (Sec. 2.3.1) lies within a small range of frequencies around ω_c. Measuring this spectrum, which contains information about the scatterers, requires clever techniques and highly monochromatic sources because of the extremely narrow spectral width. We touch briefly on quasielastic scattering in Section 3.4.

Although quasielastic scattering is *inelastic*, the qualifier inelastic often indicates *discrete* shifts of a scattered spectral line relative to an incident line, for example, Raman scattering (Sec. 5.5.1), rather than *continuous* Doppler *broadening* of an incident line because of scattering by gas molecules distributed in velocity. Inelastic scattering is distinguished from fluorescence and phosphorescence (*photoluminescence*). The distinction between the two was historically the relatively much longer duration of the latter upon abruptly cutting off the source of illumination, an inadequate criterion because of long-lived fluorescence and short-lived phosphorescence.

For an excellent discussion of photoluminescence, especially its history, see Valeur and Berberan-Santos (2011).

In atomic and nuclear physics, elastic scattering usually means that the *total* kinetic energy of projectile and target (e.g., neutron and nucleus) is conserved in a collision, which is consistent with our definition of elastic scattering of light.

1.5 Scattering, Emission, and Absorption: Theoretical Interpretation

We turn now from observations to interpretations, which require invoking theories. The essence of good physics—or so we are often told—is unification, the recognition that what were thought to be disparate phenomena are the same under the skin. To an extent, optics has been unified in that light is governed by the same equations as electromagnetic fields (Sec. 2.1). But unification has not been complete. Swords are still drawn, and blood is spilled in battles, over whether an observation is "really" caused by scattering, diffraction, interference, reflection, or refraction and whether it falls within the realm of "physical" optics or "geometrical" (unphysical?) optics or "quasioptics". Yet every diffraction pattern is an interference pattern, and diffraction, refraction, and specular reflection can be subsumed under the general heading of scattering or, more precisely, the interference of scattered waves. Lorenz (1898) understood this: "Although we have long considered light as a form of rays which interfere, are refracted and reflected at the surface of bodies following certain laws, our understanding of the luminous movement will be only elementary and fragmentary if we break up the general law of such movement into particular and separate laws of phenomena that are essentially the same." Sommerfeld (1923) concluded a note on Brewster's law (Sec. 2.10.2) with "we wish to emphasize that the phenomenon of reflection (as well as that of refraction and diffraction) is identical with that of scattering, in the classical treatment of these phenomena." This is the considered opinion of a mature Sommerfeld after many years of thinking about optics.

Separate laws are accidents of history. The light we detect or see (unless we look directly at sources of emitted light) is often scattered light regardless of what it is called. What is important is not assigning optical phenomena to pigeonholes but determining the pattern of scattered light for a given source of illumination and scatterers of specified size, shape, composition, and spatial arrangement using the simplest adequate theory. In 1927 Michelson wrote "For optical measurements two sorts of instruments are in use, to wit, lenses (or mirrors) and prisms (or gratings). The interferometer is sometimes added as a third; but because all optical phenomena depend essentially upon interference in its most general conception, there is no essential difference between them" (Michelson, 1962, p. 21). Stone (1963, p. 335) is one of the few textbook authors to recognize that "by using the word scattering in the broadest sense, one comes to a unified view of many phenomena, some of which go under different names other than scattering and are treated in practical calculations by methods that contain no explicit reference to scattering." By scattered light is sometimes meant that lateral to the direction of a nearly unidirectional beam, thereby undermining the essential unity of scattering patterns regardless of what they may be called for historical rather than physical reasons. But a beam transmitted in a palpable medium is partly composed of inseparable light scattered in the near-forward direction. Light scattered in other directions can be qualified as lateral.

Thomas Young (1802) usually is credited with the concept of interference. His Proposition VIII describes its essence: “When two Undulations from different Origins, coincide, either perfectly or nearly in Direction, their joint effect is a Combination of the Motions belonging to each.” He describes two *extremes*, constructive and destructive, and says that the intermediate “cannot be determined without further data.” But there is no evidence in this or a later paper (1804) that he ever performed the optical two-slit interference experiment often bearing his name (Rothman, 2003, Ch. 2). He implies that interference is a principle in his encyclopedic course of lectures (1807, p. 467) accompanied by the diagram of a two-source interference pattern (his Fig. 267) reproduced in countless textbooks, engendering the belief that Young *performed* a two-slit *optical* experiment. In an excellent article by Cohen (1940) about interference he refers to Young’s “famous experiment in which light from a single source is divided by means of two adjacent pinholes into two identical beams that were made to interfere.” But this is unsupported even though Cohen is otherwise scrupulous about quotations and sources. In a critical analysis of Young’s ideas, Worrall (1976) claims, followed by supporting arguments, that “there are… sufficiently many suspicious aspects of Young’s account of the two slit case to support the belief that he never performed the experiment!” Young gave insufficient details—slit width, slit separation, angular subtense of the source at the slit screen, distance to the observation screen, fringe separation—to enable others to repeat without guidance what would have been a difficult experiment in the early 19th century, and there is no evidence that anyone failed to or did repeat it until Fresnel. Worrall’s aim is to give a carefully considered rather than what might be called the Hollywood account of why Young’s wave theory of light did not immediately result in wholesale abandonment of the corpuscular theory. See also Cantor (1984) for a nuanced view. Unlike many of their dogmatic followers, Young was not a die-hard partisan of *a* wave theory nor was Newton a die-hard partisan of *a* corpuscular theory, hardly a surprise given their undisputed brilliance. Both were aware of the strengths and weaknesses of the two kinds of theories, which evolved over time and continue to do so.

The apparent puzzle of two light beams crossing without impeding each other can be explained by arguing that, because photons carry no charge, photon–photon scattering is exceedingly weak and not yet observed at visible frequencies. An alternative explanation is that electromagnetic fields act on charges and currents, not on each other, although they can add vectorially. The explanation was different in the 1800s. Herschel (1828, Article 543) argued that the “luminous molecules” (light corpuscles) were separated by such large distances that this “consideration frees us from all difficulties of their jostling or disturbing each other in space, and allows of infinite rays crossing at once through the same point of space without at all interfering with each other, especially when we consider the minuteness which must be attributed to them.”

According to classical electromagnetic theory the cause of scattering is that usually electrically neutral matter is composed of negative electrons and positive nuclei, the sources of electric and magnetic fields. Consider first a single molecule. The electric field of a time-harmonic incident electromagnetic wave exerts forces on positive and negative charges in the molecule, which causes them to accelerate and hence their fields to oscillate in space and time. Because the speed of light is finite, these changes take a finite time to be felt at all distances from the molecule: changes in the fields are propagated outwards as scattered electromagnetic waves (Sec. 3.1). Matter is mostly free space, appearances to the contrary. Because electrons are thought to have zero volume, the volume fraction of solids and liquids occupied by mass (atomic nuclei) is $\approx 10^{-13}$. The notion that electromagnetic waves of some frequencies cannot penetrate some objects, as if these waves were repelled, is untenable. Although you cannot poke a hole in an aluminum plate with your finger because of repulsive forces between charges in your finger and in the plate, electromagnetic waves with neither charge nor mass cannot be repelled *by* matter, although they do exert forces *on* it, which can result in scattered waves that may cancel (on average) the incident wave, a kind of negative feedback.

Although scattering is by individual molecules, the pattern of light scattered by a collection of molecules depends on their arrangement in space because phase differences between waves scattered by different molecules depend on their relative positions. Scattering by water vapor, a water droplet, a cloud of droplets, and a water film are quite different mostly because of *geometry*, not composition.

We can attribute scattering to a molecule as a *single* entity rather than to its separate charges *if* the wavelength is large compared with molecular dimensions (Crosignani et al., 1975, p. 4). This implicit assumption underlies assertions about scattering *by* molecules. At X-ray wavelengths comparable to atomic dimensions, however, the scatterers are atomic electrons (Als-Nielsen and McMorrow, 2001, p. 6).

Suppose that the molecules in an aggregation are identical in every respect, including orientation. This results in a particular scattering pattern. If these molecules are identical but anisotropic and not all oriented the same, the pattern of scattered light is different. In an aggregation of molecules of different kinds, not only are phases different because of geometry, so are amplitudes and phases of waves scattered by individual molecules because of their different composition and structure. Any change in the spatial distribution of molecules, even slight, for example, by high-frequency (≈MHz) sound waves in illuminated liquids, can result in measurable changes in patterns of scattered visible light (Debye and Sears, 1932).

The notion that an incident plane wave is "deformed" by interaction with a particle (which could be a lens) is a metaphor. An incident wave plane wave excites a *nonplanar* scattered wave. But the incident wave is unaffected, although it is coherent with and hence may interfere with the scattered light it excites.

If a beam of light illuminates an interface between pure water and air (Fig. 1.1) we may observe refracted and specularly reflected light, the laws of which were known empirically long before the nature of light and matter were better understood. These geometrical laws can be derived from continuum electromagnetic theory and extended to yield the *magnitude* of reflection and refraction, the Fresnel coefficients,

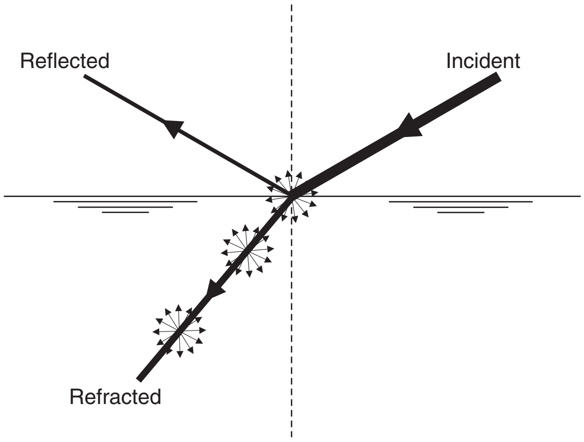

Figure 1.1 Specular reflection and transmission (refraction) because of an optically smooth boundary between optically homogeneous media (e.g., air and water). Also depicted is feeble but observable diffuse scattering because of irreducible heterogeneity on the molecular scale. The thickness of the rays depicts their relative magnitude.

for different angles of incidence and states of polarization (Sec. 2.6.2). If we look carefully, however, we may observe faint light not accounted for by continuum theory: light weakly scattered around the direction of specular reflection, called diffuse reflection and transmission in directions lateral to the plane of incidence within the illuminated medium. No matter how smooth and clean the interface, no matter how diligently we purge water of suspended matter, scattered light not invited by continuum theory crashes the party. To observe the metaphorical bending of a beam of light predicted by the law of refraction requires ignoring what it omits.

Continuum electromagnetic theory applies to hypothetical media assumed to be homogeneous with interfaces smooth on all scales. The result often is a tractable and accurate *approximate* theory adequate for describing quantitatively the dominant features of, for example, refraction and specular reflection, but not *all* the details. The alternative, rarely feasible, is to specify the position of each molecule and determine its response to excitation by an external source and by all other molecules. Continuum theory is a theory of *average* fields, which are always accompanied by *deviations*, possibly small but may be observable (Sec. 5.8). Unless stated otherwise, electromagnetic theory unqualified means continuum electromagnetic theory.

Refraction and specular reflection described by continuum electromagnetic theory are sometimes called the coherent part of the total scattered field, and everything else is the incoherent part. But there is no sharp line between these two fields, and this is a division of theories rather than of physical reality. Scattering by anything is a Gestalt. Foldy (1945) recognized this: "the connection between the so-called coherent scattering (leading to the macroscopic phenomena of refraction and specular reflection) and the incoherent scattering can only be appreciated through a consistent wave treatment of scattering from beginning to end."

Scattering by any collection of scatterers (particles or molecules) *fixed* in place is completely coherent *provided* the coherence volume of the source is larger than the collection, and hence the total field is the coherent superposition of an incident field and a scattered field, the square of which is observed with a detector that responds to power. We can average this square—but what is meant by an average? We have two possibilities: ensemble or time. By ensemble average is meant an average over many systems macroscopically but not microscopically identical. For example, any pure water droplet is unlike any other water droplet with the same diameter and at the same temperature even if we assume that their molecules don't move. If they move, we can take the time average over a single droplet. It is often assumed that these two averages are equal (the *ergodic hypothesis*). By the square of a field is meant the product of its complex representation (or its Fourier transform) with its complex conjugate (see Sec. 2.4).

To elucidate the difference between ensemble and time averages and their dependence on the coherence properties of illumination and illuminated object, consider a sheet of white paper, a coherent array of fibers fixed in place. Illuminated by sunlight or lamplight, the result is uniform brightness. But if illuminated by an expanded laser beam the result is a mottled pattern of bright and dark, called *laser speckle*, an interference pattern resulting from illumination of a coherent object by a coherent source. An average over the paper is an ensemble average, and we expect it to be equal to what is observed with lamplight or solar illumination, which in effect performs a time average. If a glass of milk is illuminated by the laser beam, laser speckle is not observed, only uniform brightness. The coherence of the waves scattered by the milk particles is effaced by random buffeting by liquid molecules, the result being a time average *as if* phase differences did not exist. With a laser beam expander, you can observe these two different patterns for yourself.

Continuum electromagnetic theory yields a good *estimate* of the mean field scattered by a droplet, the square of which can be compared with measurements. The mean field has "little physical significance" unless the relative difference between the mean of the square and the square of the mean is small. A cloud of many water droplets does not satisfy this condition, and hence is treated by a theory (radiative transfer) of the mean square field (radiance), the mean field bypassed. Any difference between what is predicted by continuum theory and what is observed or predicted by a more detailed theory can be lumped under the heading of incoherent scattering.

Reflected light not accounted for by theories in which interfaces are absolutely smooth is sometimes called "surface scattering," whereas scattered light within an illuminated object is sometimes called "volume scattering." But light scattered in *all* directions is a consequence of scattering by *all* the molecules of the object. See Section 2.8 for the scattering interpretation of specular reflection. An incident wave does not "bounce" off a surface, a stale and often misleading metaphor, rather it excites a highly directional scattering pattern by a huge phased array of molecular-scale antennas.

The term molecular scattering is sometimes applied to the incoherent part of the scattering, which is slightly jarring because it implies the existence of *non*molecular scattering. And yet the same molecules that are the source of refracted and specularly reflected light are the source of the rest of the scattered light. How could this be otherwise? Molecules scatter light, so again the distinction is not between fundamental processes but between different theories and how finely we probe. For example, lenses can be designed without explicitly considering their molecular composition or agonizing over the nature of light. By molecular scattering of light usually is meant light unaccounted for by continuum electromagnetic theory (Sec. 5.8).

Molecules can translate, rotate, and vibrate resulting in temporal variations in light scattered in each direction by a collection of molecules. They can move independently of each other or in completely or partially correlated groups. No matter what we call temporal variations in scattered light, their underlying physical origin is the same.

There are many theories of light scattering by particles (or anything) but only one cause. We are careful to distinguish between theories and causes, which avoids much pointless arguing. The same physical phenomenon often can be described to the same degree of approximation by different theories with different names and that superficially appear different. None of these theories can be said to be *the* cause of the phenomenon.

The basic theory adopted in this book is classical continuum electromagnetic theory, although we make brief excursions into quantum theory in Chapter 9. In principle, scattering by a water droplet can be treated by quantum electrodynamics, which is as silly as saying that in principle quantum electrodynamics can forecast the weather.

Classically, emission by a particle can be spontaneous, the result solely of the random but bounded nonuniform motions of its constituent charges. If accelerating charges radiate, and matter is composed of charges moving back and forth, emission is a necessary consequence, and is always accompanied by a change in the state of the emitter. Absorption is the result of incident fields doing work on these moving charges, which results in a change in the state of the absorber. This interpretation of emission and absorption may carry over into the quantum domain. For example, Rothman and Sudarshan (1998, p. 285) aver that "what we call the creation or absorption of a photon is just a transition from one state to another. We write down the initial state and final state and see that the state of the system has changed. We associate this change of state with the creation or absorption of a photon. But it is just a transition. There is no mechanism in the theory for creating or destroying anything." For a classical interpretation of *stimulated emission* see Section 9.1.1.

Absorption and scattering are not mutually exclusive. For brevity, we often refer only to scattering, but always have in mind absorption as well. A particle (or molecule) that only scatters or only absorbs is an idealization.

1.6 Physics of Scattering by a Single Particle

Molecules are hidden offstage in continuum electromagnetic theory. In any scattering problem tackled with this theory the only scale parameters are characteristic linear dimensions of illuminated objects. These parameters are sufficient to determine *gross* phase differences between scattered waves but not the small contributions to phase differences from heterogeneity on the molecular scale.

All particles, regardless of composition, are penetrable by electromagnetic waves, which *must* excite charges within them to scatter. A particle may be impenetrable to *acoustic* waves, which do not propagate in free space, whereas electromagnetic waves do.

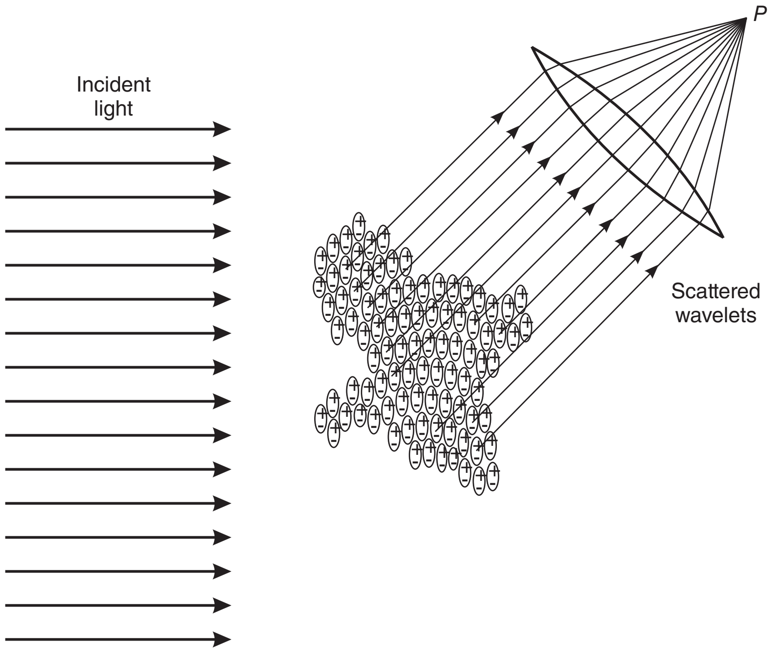

Figure 1.2 The total scattered field at P is the coherent sum of all the wavelets scattered by the dipoles into which the particle is imagined to be subdivided.

It is not feasible to calculate scattering by, say, a water droplet of diameter 10 μm by adding the small amplitude wavelets scattered by each of its 10^{13} interacting molecules, if only because their exact positions are not known. But we can imagine this to acquire understanding. Scattering by a particle considered as an ensemble of identical microscopic (molecular) dipoles is depicted *qualitatively* in Figure 1.2; (7.64) is its *mathematical* form. Each dipole excited to oscillate with amplitude smaller than atomic dimensions, at the frequency of the incident primary wave scatters a secondary wavelet in every direction. Because 10^{13} dipoles interact with each other, scattering by the droplet is inherently, but not necessarily explicitly, an electromagnetic *many-body problem* (Mattuck, 1976, p. 1). For an *approximate* way to transform an array of a huge number of *microscopic* dipoles into a much smaller number of *macroscopic* dipoles, see Section 8.11.1.

The total *internal* (or *local*) field exciting each dipole is the superposition of the incident field and the fields of the other dipoles in the particle. The total *external* scattered field in any direction at a distant point P is the superposition of many wavelets with phases depending on the positions of the dipoles that created them and the phase of the field exciting the dipoles; the amplitude of each wavelet is proportional to the local field that excites it. Phase differences between wavelets change with scattering direction, in general, and hence so does the total scattered wave. But if a particle is small compared with the wavelength, all the wavelets are approximately in phase, and we expect a smooth variation of scattering with direction. We also expect the total scattered *power* to increase as the square of the number of dipoles (square of the total scattered field), and hence as the particle volume squared. As the particle size relative to the wavelength increases, the possibilities for enhancement *and* cancellation of scattered wavelets increase, and hence the more the scattering pattern varies with direction. For the same reason, total scattered power must increase less steeply than volume squared. If we change the overall *shape* of a particle by rearranging a fixed number of dipoles, all the phase differences are changed, and hence so is the scattering pattern. These expectations based on physical reasoning are supported by detailed calculations for specific particles. We show in Section 3.4 that by going a small step further we can obtain quantitative, although approximate, results by considering only *two* dipoles.

Phase differences between scattered wavelets depend on geometry—scattering direction, particle size, and shape—whereas the amplitude and phase (relative to the incident wave) of each wavelet depend on particle composition. To fully understand scattering and absorption by particles, we therefore need to know how bulk matter responds to excitation by fields of different frequencies (see Chs. 9 and 10).

An assumption implicit in Figure 1.2 is that the particle is in free space. If it is in a material medium, to claim that, unlike the particle, this medium is not composed of a great many molecules is inconsistent. Molecules *outside* the particle also scatter waves that superpose on each other throughout the medium and the particle and on the waves from molecules *inside* the particle to yield the scattering pattern. This is the fundamental reason why scattering depends on the composition of the particle *and* of the surrounding palpable medium, which is predicted by continuum electromagnetic theory without explaining why. Except at the surface of a particle, its molecules are surrounded only by like molecules. Although molecules from the surrounding medium do not physically penetrate the particle, they change the total scattered field because of the waves they scatter. This is rarely stated, but it follows from the microscopic interpretation of scattering. In macroscopic theories, the assumption that both particle and surrounding medium are continuous is consistent. Consistency also requires that in microscopic theories the discreteness of both is acknowledged. What can be done in practice, however, is to consider a particle to be composed of dipoles each of which is suspended in a surrounding continuous medium (Ch. 6).

Computational methods and approximate theories for calculating scattering by particles sometimes obscure the underlying physics but are equivalent to adding scattered wavelets, generally with different phases, by huge numbers of dipoles in free space. Scattering diagrams are interference patterns even if described by theories in which the wave nature of light is not explicit. For certain classes of particles, scattered fields may be calculated by subdividing particles into noninteracting dipolar regions (not molecules) and adding the wavelets they scatter (Ch. 6).

Although we often say that the incident (or primary) field excites a particle to scatter, which is true in a sense, the exciting field is the *net* field at every point *within* a particle, which also determines the electromagnetic energy absorbed (Sec. 3.4). The incident field excites a secondary field that superposes on the incident field. These fundamental facts are often lost sight of because of the way rigorous and approximate theories are formulated. For example, there is plenty of multiple scattering within a single particle, but not explicit in theories of scattering based on continuum electromagnetic theory (Sec. 2.1).

Democritus (460–370 BC) got it right long ago: "Nothing exists except atoms and empty space; everything else is opinion." The theories in this book are opinions, not dogma.

1.7 Direct and Inverse Problems

There are two general classes of problems in the theory of scattering of electromagnetic waves by small particles:

The Direct (or Forward) Problem. For a particle of known size, shape, and composition, illuminated by a beam with a given irradiance, polarization, and frequency, determine the scattered field everywhere and the rate of absorption. This is the "easy" problem: describe the tracks of a given dragon (Fig. 1.3a).

The Inverse (or Backward) Problem. By a suitable analysis of measured scattered radiation for a range of frequencies of the incident radiation and scattering directions, describe the particle or particles responsible for the scattering. Although we cannot measure absorbed radiation directly, we may be able to measure its consequences (e.g., temperature changes), which gives additional information. This is the "hard" problem: describe a dragon by examining its tracks (Fig. 1.3b).

The two classes of problems are not independent. Solving the inverse or diagnostic problem (effect → cause) is not possible without solving the direct or predictive problem (cause → effect), but not conversely.

The inverse problem, the kind to which Sherlock Holmes might have directed his reasoning powers, often is of greatest interest. For example, although interstellar dust is known to be abundant in our galaxy, the exact composition of this dust is

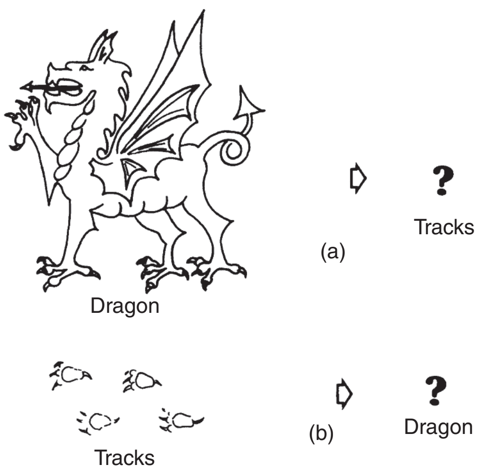

Figure 1.3 (a) The direct problem: Describe the tracks of a given dragon. (b) The inverse problem: Describe a dragon from its tracks.

not known despite considerable efforts to determine it. The only way to infer its composition is to analyze the light of various frequencies it transmits without being scattered or absorbed or, less commonly, the light it scatters in various directions. In laboratory investigations or in industrial processes, light scattering is often used to infer the size of particles of known shape and composition. Radar backscattering can be used to discriminate between rain and hail. In these applications the goal is to describe a dragon or flock of dragons by detailed examination of their tracks.

To understand why we call this the "hard" problem, to specify a particle uniquely requires the vector amplitude and phase of the field scattered in every direction or the internal field (Hart and Gray, 1964). The internal field usually is not accessible to direct measurement, although under certain conditions, unlikely to be met except in the laboratory, it can be *approximated* by the incident field (Ch. 6). Even for this special case, however, the amplitude and phase of the scattered field are required, obtainable in principle but rarely in practice. Because measurements available for analysis usually are of scattered irradiances for a limited range of frequencies and directions, we almost always face the task of trying to describe a particle or a collection of particles by analyzing a less than ideal set of data. But this is not necessarily cause for despair. Often, supplementary information about particles (*prior knowledge*), obtained by means other than light scattering, may be sufficient for them to be described. Given that we rarely have enough information, we should be reluctant to discard information that is or might become available. One source of information sometimes neglected is the polarization state of scattered light, which we treat in detail in the following chapters.

The two basic methods of inverting scattering data to obtain particle characteristics are nicely illustrated by an analogy attributed to van de Hulst. You can buy a suit in two ways. Try on suits off the rack in a clothing store until you find the best fit. According to Baltes (1978, p. 2), "The notion of inverse problem in the strict sense is usually understood to exclude such fitting procedures," for example, using lookup tables of calculations to match measurements. The alternative is to dig deeper into your pockets and have a set of measurements made—waist, inseam, arm length, etc.—to be transformed into a suit by a skilled tailor. This is analogous to a formal mathematical inversion method in which measurements (with errors) are fed into an algorithm that spits out particle characteristics. But for many applications, especially inverting measurements made on distributions of remote particles, trial and error may be the only feasible method.

Both inversion schemes face problems of the *sensitivity* of measurements (always with errors) to the particle characteristic of interest (see, e.g., Sec. 13.8.2) and *uniqueness*. Mathematical inversion schemes also must be *stable*: small changes in input must correspond to small changes in output.

Even if we restrict ourselves to a single uniform particle with a well-defined shape, it is unrealistic to expect to be able to infer its characteristics—size, shape, composition—solely from light scattering measurements. A homogeneous sphere of *known* composition has only one degree of freedom, its diameter, and the direct problem for arbitrary diameter and (uniform) composition is soluble. A spheroid has two degrees of freedom, but the direct problem is not so readily solved.

An ellipsoid has three degrees of freedom, and the direct problem is even more complicated. When we turn to irregular particles—even if we could agree on how to define them—the number of degrees of freedom is indefinite, and we lack a rigorous solution to the direct problem. Only for single, homogeneous spheres of known composition, can we hope to obtain their size. For any other particles, the most we can expect is *estimates* of size, shape, and composition.

Annotated References and Suggestions for Further Reading

Nearly a century after Zsigmondy saw flashes of light from tiny gold particles below the limit of resolution, Neuhauser et al. (1980) saw light from a single barium ion, a refutation of Mach's assertion (1960, p. 588) that "atoms cannot be perceived by the senses."

For a good history of the ultramicroscope, see Cahan (1996).

Tyndall observed scattering by ultramicroscopic particles a few decades before Zsigmondy, noting that a "concentrated beam reveals them collectively long after the microscope has ceased to distinguish them individually" (Tyndall, 1888, p. 77).

For an excellent treatment of microscopy, including translations of papers by Abbe and Helmholtz, see Masters (2020).

Despite his fame for precision interferometry, Michelson (1962, p. 10) criticized the term "interference" although resigned to it: "The term [interference] is not very well chosen, for in fact each [wave] train produces its own effect quite independently of the other, but it has been in use so long that it would not seem wise to alter it." Even Lord Rayleigh expressed sentiments similar to Michelson's in his 1888 *Encyclopedia Britannica* article "Wave theory of Light": "the word is not very appropriate, the vibrations being simply *superposed* with as little interference as could be imagined." Waves can "interfere" in more than one sense: interference often denotes unwanted signals in radio communication and radar. For a history of interference of light, see Kipnis (1991), who notes (p. 18) that "[in] general, physicists are very vague about the meaning of 'interference.'"

For a brief, amusing account of a "long, bitter polemic" between Lommel and Exner over whether a phenomenon was caused by diffraction or scattering, see De Witte (1967, footnote 4).

See Magie (1963, pp. 294–298) for a translated excerpt from Grimaldi's posthumous 545-page *Physico-mathesis de Lumine, Coloribus et Iridae* (1665), in which he records observations of spatial patterns, now attributed to diffraction (see Sec. 4.4.3), not in accord with then-current views about shadows.

For a definition of the term "quasioptics," its scope, and history, see Goldsmith (1998, Ch. 1).

Among the books devoted entirely to scattering by particles, that by Shifrin (1951) most closely resembles ours because it discusses optical properties of bulk matter. Although the translation from Russian is in places awkward, the terminology is sometimes confusing, symbols are difficult to decipher, and the figures are of low quality, this book is worth reading for its insights and to set the historical record straight. Shifrin and other Russian scientists are not always given due credit for results subsequently published by North American and European scientists.

The two books that have influenced us most are those by van de Hulst (1957) and by Kerker (1969), to whom we are indebted. Deirmendjian (1969) emphasizes scattering by polydispersions. Papers in the collection edited by Bowman et al. (1969) are restricted to scattering by infinite or bounded objects composed of materials either with a relative refractive index close to 1 or infinite. The treatise by Mishchenko et al. (2002) in some ways parallels ours, but they favor the integral rather than the differential formulation of scattering and present many more computational results than we do. This wealth of computations and a thorough treatment of T-matrices (see Sec. 8.11.2) are the book's strongest points. Scattering of visible light by nonmetallic particles, including those of biological origin (bacteria, phytoplankton), suspended in seawater and freshwater is treated in detail by Jonasz and Fournier (2007). They emphasize measurements and include extensive appendices on size distributions and nephelometer design (Sec. 13.2.1).

The "main subject" of Kokhanovsky's book (2004a, p. 1) is "calculation of the characteristics of a light field in disperse media." To that end, two chapters on single particles are followed by chapters on radiative transfer theory and applications in geophysics and biology.

Although somewhat dated, a collection of conference papers edited by Kerker (1963), about one-third of which is devoted to single-scattering by particles, is worth perusing. For a handy two-volume compendium of selected papers on light scattering by particles, see Kerker (1988).

About one-sixth of a treatise on scattering theory by Newton (1982) is devoted to scattering of electromagnetic waves by particles.

Bayvel and Jones (1981) treat in about 70 pages what we discuss in more detail in eight chapters. But they go beyond single particles to discuss theories and methods for treating multiple (incoherent) scattering, methods for inferring size distributions, and finally give many applications such as determining particle size distributions in, for example, steam turbines, sprays, and flames.

Two volumes by Ishimaru (1978a, 1978b) cover many topics in electromagnetic and acoustic scattering. Sources are not necessarily steady in time or laterally infinite nor are the scatterers necessarily stationary. Only one chapter covers absorption and scattering by single particles. The first volume (1978a) is mostly about radiative transfer theory, including various approximations to the equation of transfer. The second volume (1978b) is mostly about coherent multiple scattering, and includes random continuous media (e.g., turbulent air) and scattering because of surface roughness. The final chapter is on remote sensing. He calls "analytical theory" or "multiple scattering theory" what we call multiple scattering of waves or coherent multiple scattering. Transport theory in his lexicon is based on an equation of transfer derived without explicitly invoking waves and their coherence properties.

The integro-differential equation for the spatial and directional distribution of radiation in an incoherently scattering medium is essentially the Boltzmann transport equation for the distribution of molecular velocities in a gas sufficiently dilute that only binary collisions need be accounted for (see, e.g., Reif, 1965, p. 525). Molecular energies can change by collisions whereas photon energies cannot (by elastic scattering), and although gas molecules are not destroyed or created (absent chemical reactions), photons can be imagined to be absorbed and emitted. The equation of transport for neutrons is even more like the radiative transfer equation (Glasstone and Edlund, 1952, p. 388).

Mishchenko (2013) recounts the history of radiative transfer and criticizes the phenomenological theory.

For a derivation of the probability distribution of nearest-neighbor distances between randomly distributed particles and its mean, see Chandrasekhar (1943, App. VII).

To understand coherent multiple scattering, the best place to begin is Foldy's 1945 paper in which he considers the simplest problem: elastic, isotropic scattering of scalar waves by a nonabsorbing, nonemitting medium (for an "appreciation" of Foldy, see Taylor and Fickinger, 2007). If you crave more, you can turn to Lax (1951), who removes these restrictions and includes anisotropic scattering, inelastic scattering, and scatterers "randomly, partially, or completely ordered." He does not restrict himself to electromagnetic waves. Twersky (1962) treats in detail multiple scattering of waves by a regular linear array of identical scatterers (grating) as well as by rough surfaces and scatterers randomly distributed in space. For a mathematical treatise on coherent multiple scattering, see Martin (2006). Sources of illumination are time–harmonic, and the scatterers, finite in number and bounded, are fixed in regular, as opposed to random, positions. He casts a wide net that captures acoustic, electromagnetic, elastodynamic, and hydrodynamic waves.

Although Chandrasekhar's (1960) treatise on radiative transfer, first published in 1950, is considered the classic work on the subject, it is not recommended for neophytes. One of his major contributions was to incorporate polarization in radiative transfer theory by way of the Stokes parameters (Sec. 2.10.1). van de Hulst's (1980) two-volume treatise is "aimed at the nonspecialist" and includes chapters on astronomical and atmospheric applications. In contrast with the treatises by Chandrasekhar and by van de Hulst, in which electromagnetic fields are absent, Mishchenko et al. (2006) begin with the Maxwell equations, discuss both coherent and incoherent multiple scattering, and include a chapter on coherent backscattering. They pay special attention to polarization. Their aim is to firmly ground phenomenological radiative transfer theory in electromagnetic theory rather than present a catalogue of methods for solving the equation of radiative transfer derived heuristically. They limit their treatment to "randomly and sparsely positioned particles." An advanced textbook by Thomas and Stamnes (1999) devotes several chapters to solution methods for the equation of radiative transfer. The applications are to radiation in the atmosphere and ocean. For a treatment of incoherent multiple scattering aimed at explaining the greatest number of observations with the least amount of mathematics, see Bohren (1987a); for an even simpler exposition, see Bohren and Clothiaux (2006, Ch. 5). Seminal papers on radiative transfer by Schuster, Schwarzschild, Eddington, Rosseland, and Milne are gathered in a compendium edited by Menzel (1966).

Review chapters on absorption and scattering by single particles are included in *Light Scattering Reviews*, edited by Kokhanovsky and published by Springer (Berlin). As of 2014, nine volumes had appeared, the first in 2006. They also include reviews of multiple scattering theories and techniques, remote sensing, and applications to Earth's atmosphere, oceans, and snow and ice cover.

For a good, concise discussion of coherence, see Winter et al. (n.d.).

In his paper on scattering of charged particles by atoms, which engendered nuclear physics, Rutherford (1911) used *compound scattering* to mean multiple scattering. This term appears to have died with him, although compound elastic scattering with a different meaning is used in nuclear physics.

The classical treatise on molecular scattering of light is by Cabannes and Rocard (1929). Among its successors are treatises by Bhagavantam (1942), about half of which is devoted to Raman scattering and its application in various branches of chemistry, and by Fabelinskii (1968).

For quasielastic scattering see Chu (1991), also the collection of papers compiled by Chu (1990).

Inverse problems in scattering are discussed in collections of papers edited by Baltes (1978, 1980). The journal *Inverse Problems* occasionally publishes papers on inversion of electromagnetic radiation scattered by particles.

There are similarities (and differences) between absorption and scattering of electromagnetic waves and of acoustic waves by suspensions of particles. For an excellent treatment of suspension acoustics see Temkin's (2005) Figs. 1.1.1–1.1.3 and Chapter 9.

Books and papers relevant to specific chapters are cited at their ends.