# Chapter 1

# Introduction

Natural and man-made environments provide countless examples of diverse scattering media composed of particles. The varying complexity of these media suggests multiple ways of using electromagnetic scattering for particle characterization and gives rise to a distinctive hierarchy of theoretical models that can be used to simulate specific remote-sensing or laboratory measurements. Hence the objective of this introductory chapter is to present a simple classification of scattering problems involving small particles and to briefly outline solution approaches described in detail in later chapters.

## 1.1 Electromagnetic scattering by a fixed finite object

A parallel monochromatic beam of light propagates in a vacuum without a change in its intensity or polarization state. However, inserting an object into the beam (see Fig. 1.1.1) causes several distinct effects. First, the object extracts some of the incident energy and spreads it in all directions at the frequency of the incident beam. This phenomenon is called elastic *scattering* and, in general, gives rise to light with a polarization state different from that of the incident beam. Second, the object may convert some of the energy contained in the beam into other forms of energy such as heat. This phenomenon is called *absorption*. The energy contained in the incident beam is accordingly reduced by the amount equal to the sum of the scattered and absorbed energy. This reduction is called *extinction*. The extinction rates for different polarization components of the incident beam can be different, which is called *dichroism*.

In electromagnetic terms, the parallel monochromatic beam of light is represented by a harmonically oscillating plane electromagnetic wave. The latter propagates in a

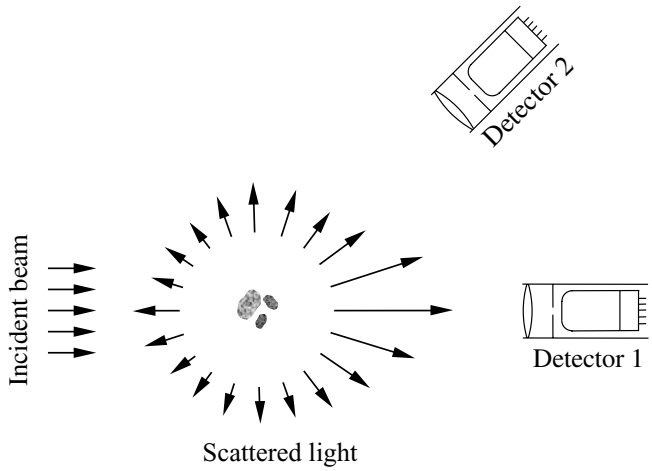


**Figure 1.1.1.** Scattering by a fixed finite object. In this case the object consists of three disjoint, heterogeneous, stationary bodies.

vacuum without a change in its intensity or polarization state (see Fig. 1.1.2(a)). However, the presence of a finite object, as illustrated in Fig. 1.1.2(b), changes both the electric, $\boldsymbol{\mathcal{E}}$, and the magnetic, $\boldsymbol{\mathcal{H}}$, field that would otherwise exist in an unbounded homogeneous space. The difference between the total fields in the presence of the object, $\boldsymbol{\mathcal{E}}(\mathbf{r},t)$ and $\boldsymbol{\mathcal{H}}(\mathbf{r},t)$, and the original fields that would exist in the absence of the object, $\boldsymbol{\mathcal{E}}^{\rm inc}(\mathbf{r},t)$ and $\boldsymbol{\mathcal{H}}^{\rm inc}(\mathbf{r},t)$, can be thought of as the fields scattered by the object, $\boldsymbol{\mathcal{E}}^{\rm sca}(\mathbf{r},t)$ and $\boldsymbol{\mathcal{H}}^{\rm sca}(\mathbf{r},t)$, where $\mathbf{r}$ is the position (radius) vector and $t$ is time (Fig. 1.1.2(b)). In other words, the total electric and magnetic fields in the presence of the object are equal to vector sums of the respective incident (original) and scattered fields:

$$\boldsymbol{\mathcal{E}}(\mathbf{r},t) = \boldsymbol{\mathcal{E}}^{\rm inc}(\mathbf{r},t) + \boldsymbol{\mathcal{E}}^{\rm sca}(\mathbf{r},t), \tag{1.1.1}$$

$$\boldsymbol{\mathcal{H}}(\mathbf{r},t) = \boldsymbol{\mathcal{H}}^{\rm inc}(\mathbf{r},t) + \boldsymbol{\mathcal{H}}^{\rm sca}(\mathbf{r},t). \tag{1.1.2}$$

The origin of the scattered electromagnetic field can be understood by recalling that in terms of microscopic electrodynamics, the object is an aggregation of a large number of discrete elementary electric charges. The oscillating electromagnetic field of the incident wave excites these charges to vibrate with the same frequency and thereby radiate secondary electromagnetic waves. The superposition of all the secondary waves gives the total elastically scattered field. If the charges do not oscillate exactly in phase or exactly in anti-phase with the incident field then there is dissipation of electromagnetic energy into the object. This means that the object is absorbing and scatters less total energy than it extracts from the incident wave.

Electromagnetic scattering is an exceedingly complex phenomenon because a secondary wave generated by a vibrating charge also stimulates vibrations of all other charges forming the object and thus modifies their respective secondary waves. As a

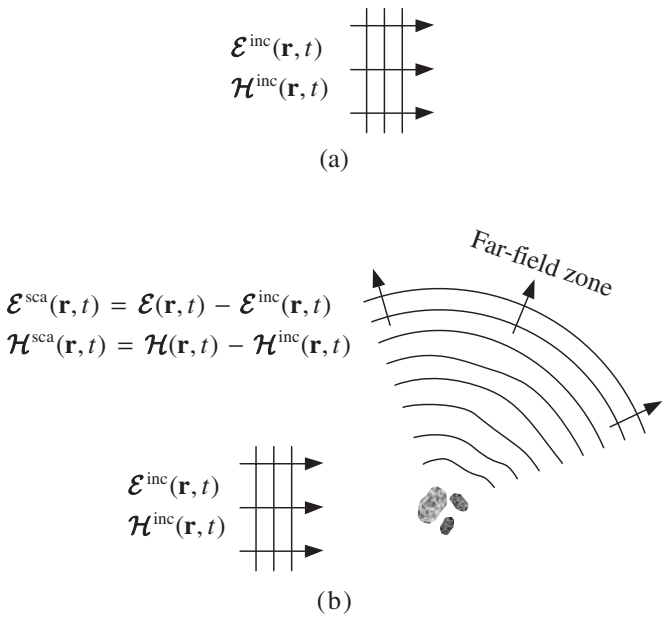


**Figure 1.1.2.** Schematic representation of the electromagnetic scattering problem.

result, all the secondary waves become interdependent. Furthermore, the computation of the total scattered field by superposing the secondary waves must take account of their phase differences, which change every time the incidence and/or the scattering direction is changed. Therefore, the total scattered field depends on the way the charges are arranged to form the object with respect to the incidence and scattering directions.

Since the number of elementary charges forming an object can be extremely large, solving the scattering problem directly by computing and superposing all secondary waves is impracticable even with the aid of modern computers. Fortunately, the scattering problem can also be solved using the concepts of macroscopic electromagnetics, which treat the large collection of charges as one or several macroscopic bodies with a specific distribution of the refractive index. Consequently, the scattered field can be computed by solving the Maxwell equations for the macroscopic electromagnetic field subject to appropriate boundary conditions. It is this approach that forms the basis of the modern theory of electromagnetic scattering by macroscopic objects.

To simplify the solution of the scattering problem, we will adhere throughout the book to the following five well-defined restrictions:

1. We will always assume that the unbounded host medium surrounding the scattering object is homogeneous, linear, isotropic, and nonabsorbing.

2. We will always assume that the scattering object is illuminated by either:

   (i) a time-harmonic plane electromagnetic wave given, in the complex-field

representation, by

$$\begin{cases} \mathbf{E}(\mathbf{r},t) = \mathbf{E}_0 \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{r} - \mathrm{i}\omega t), \\ \mathbf{H}(\mathbf{r},t) = \mathbf{H}_0 \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{r} - \mathrm{i}\omega t), \end{cases} \quad \mathbf{r} \in \Re^3, \tag{1.1.3}$$

with constant amplitudes $\mathbf{E}_0$ and $\mathbf{H}_0$, where $\omega$ is the angular frequency, $\mathbf{k}$ is the real-valued wave vector, $\mathrm{i} = (-1)^{1/2}$, and $\Re^3$ denotes the entire three-dimensional space, or

(ii) a quasi-monochromatic parallel beam of light given by

$$\begin{cases} \mathbf{E}(\mathbf{r},t) = \mathbf{E}_0(t) \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{r} - \mathrm{i}\omega t), \\ \mathbf{H}(\mathbf{r},t) = \mathbf{H}_0(t) \exp(\mathrm{i}\mathbf{k}\cdot\mathbf{r} - \mathrm{i}\omega t), \end{cases} \quad \mathbf{r} \in \Re^3, \tag{1.1.4}$$

where fluctuations in time of the complex amplitudes of the electric and magnetic fields, $\mathbf{E}_0(t)$ and $\mathbf{H}_0(t)$, around their respective mean values occur much more slowly than the harmonic oscillations of the time factor $\exp(-\mathrm{i}\omega t)$.

This restriction excludes other types of illumination such as a focused laser beam of finite lateral extent or a pulsed beam.

3. We will exclude nonlinear optics effects by assuming that the conductivity, permeability, and electric susceptibility of both the scattering object and the surrounding medium are independent of the electric and magnetic fields.

4. We will assume that electromagnetic scattering occurs without frequency redistribution, i.e., the scattered light has the same frequency as the incident light. This restriction excludes inelastic scattering phenomena such as Raman and Brillouin scattering and fluorescence. It also excludes the specific consideration of the small Doppler shift of frequency of the scattered light relative to that of the incident light due the movement of the scatterer with respect to the source of illumination.

5. We will largely exclude from consideration the phenomenon of thermal emission. The latter is caused by electron transitions from one energy level to a lower level in macroscopic bodies with absolute temperature different from zero. A macroscopic object is a complex system of molecules with a large number of degrees of freedom. Therefore, many different electron transitions produce spectral emission lines so closely spaced that the resulting radiation spectrum becomes effectively continuous and includes emitted energy at all frequencies. By neglecting thermal emission, we will implicitly assume that the temperature of the object is low enough that the intensity of the emitted radiation at the frequency of the incident light is much smaller that the elastically scattered intensity. This assumption is usually valid for objects at room or lower temperature and for near-infrared and shorter wavelengths.

The theoretical and numerical techniques for computing the electromagnetic field elastically scattered by a finite fixed object composed of one or several physical bodies are many and are reviewed thoroughly in Mishchenko *et al.* (2000a), MTL, and Kahnert (2003). Since all of these techniques have certain limitations in terms of the object morphology and object size relative to the incident wavelength, a practitioner should analyze carefully the relative strengths and weaknesses of the available solution techniques before attempting to address the specific problem in hand.

## 1.2 Actual observables

Because of high frequency of time-harmonic oscillations, traditional optical instruments cannot measure the electric and magnetic fields associated with the incident and scattered waves. Indeed, accumulating and averaging a signal proportional to the electric or the magnetic field over a time interval long compared with the period of oscillations would result in a zero net result:

$$\frac{1}{T}\int_{t}^{t+T} \mathrm{d}t' \exp(-\mathrm{i}\omega t') \underset{T \gg 2\pi/\omega}{=} 0. \tag{1.2.1}$$

Therefore, the majority of optical instruments measure quantities which have the dimension of energy flux and are defined in such a way that the time-harmonic factor $\exp(-\mathrm{i}\omega t)$ vanishes upon multiplication by its complex-conjugate counterpart: $\exp(-\mathrm{i}\omega t)[\exp(-\mathrm{i}\omega t)]^* \equiv 1$. This means that in order to make the theory applicable to analyses of actual optical observations, the scattering process must be characterized in terms of carefully chosen derivative quantities that can be measured directly. This explains why the concept of an actual *observable* is central to the discipline of light scattering by particles.

Although one can always define the magnitude and the direction of the electromagnetic energy flux at any point in space in terms of the Poynting vector, the latter carries no information about the polarization state of the incident and scattered fields. The conventional approach to ameliorate this problem dates back to Sir George Gabriel Stokes. He proposed using four real-valued quantities which have the dimension of monochromatic energy flux and fully characterize a *transverse* electromagnetic wave[1] inasmuch as it is subject to practical optical analysis (Stokes, 1852). These quantities, called the Stokes parameters, form the so-called four-component Stokes column vector and carry information about both the intensity and the polarization state of the wave.

In the so-called far-field zone of a fixed object, the propagation of the scattered electromagnetic wave is away from the object (Fig. 1.1.2(b)). Furthermore, the elec-

[1] By definition, the electric and magnetic field vectors of a transverse electromagnetic wave vibrate in the plane perpendicular to the propagation direction.

tric and magnetic field vectors vibrate in the plane perpendicular to the propagation direction and their amplitudes decay inversely with distance from the object. The tranversality of both the incident plane wave and the scattered spherical wave allows one to define the corresponding sets of Stokes parameters and to describe the response of a well-collimated polarization-sensitive detector of light in terms of the $4\times4$ so-called phase and extinction matrices. Specifically, detector 2 in Fig. 1.1.1 collects only the scattered light, and its response is fully characterized by the product of the phase matrix and the Stokes column vector of the incident wave. Thus the phase matrix realizes the transformation of the Stokes parameters of the incident wave into the Stokes parameters of the scattered wave. The response of detector 1 consists of three parts:

1. The one due to the incident light.
2. The one due to the forward-scattered light.
3. The one due to the interference of the incident wave and the wave scattered by the object in the exact forward direction.

The third part is described by minus the product of the extinction matrix and the Stokes vector of the incident wave and accounts for both the total attenuation of the detector signal due to extinction of light by the object and the effect of dichroism.

The phase and extinction matrices depend on object characteristics such as size, shape, refractive index, and orientation and can be readily computed provided that the scattered field is known from the solution of the Maxwell equations.

The main convenience of the far-field approximation is that it allows one to treat the object essentially as a point source of scattered radiation. However, the criteria defining the far-field zone are rather stringent and are often violated in practice. A good example is remote sensing of water clouds in the terrestrial atmosphere using detectors of electromagnetic radiation mounted on aircraft or satellite platforms. Such detectors typically measure radiation coming from a small part of a cloud and do not "perceive" the entire cloud as a single point-like scatterer. Furthermore, the notion of the far-field zone of the cloud becomes completely meaningless if a detector is placed inside the cloud. It is thus clear that to characterize the response of such "near-field" detectors one must define quantities other than the Stokes parameters and the extinction and phase matrices. Still the actual observables must be defined in such a way that they can be measured by an optical device ultimately recording the flux of electromagnetic energy.

## 1.3 Foldy–Lax equations

Many theoretical techniques based on directly solving the differential Maxwell equations or their integral counterparts are applicable to an arbitrary fixed finite object, be it a single physical body or a cluster consisting of several distinct components, either

touching or spatially separated. These techniques are based on treating the object as a single scatterer and yield the total scattered electric and magnetic fields. However, if the object is a multi-particle cluster then it is often convenient to represent the total scattered field as a vector superposition of partial fields scattered by the individual cluster components. This means, for example, that the total electric field at a point **r** is written as follows:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}^{\rm inc}(\mathbf{r},t) + \sum_{i=1}^{N} \mathbf{E}_i^{\rm sca}(\mathbf{r},t), \qquad \mathbf{r} \in \Re^3, \tag{1.3.1}$$

where $N$ is the number of particles in the cluster and $\mathbf{E}_i^{\rm sca}(\mathbf{r},t)$ is the $i$th partial scattered electric field. The total magnetic field is given by a similar expression. The partial scattered fields can be found by solving vector so-called Foldy–Lax equations which follow directly from the volume integral equation counterpart of the Maxwell equations and are exact. By iterating the Foldy–Lax equations, one can derive an order-of-scattering expansion of the scattered field which, in combination with statistical averaging, forms the basis of the modern theory of multiple scattering by random particle ensembles.

## 1.4 Dynamic and static scattering by random groups of particles

Solving the Maxwell equations yields the field scattered by a fixed object. This approach can be used directly in analyses of microwave analog measurements (e.g., Gustafson, 2000; Section 8.2 of MTL), in which the scattering object is held fixed relative to the source of electromagnetic radiation during the measurement cycle. However, it is inapplicable in the majority of laboratory and remote-sensing observations. Even if the scattering object is a single microparticle trapped inside an electrostatic or optical levitator (e.g., Chapter 2 of Davis and Schweiger, 2002), it rapidly changes its position and orientation during the time necessary to take a measurement. Furthermore, one often encounters situations in which light is scattered by a very large group of particles forming a constantly varying spatial configuration. A typical example is a cloud of water droplets or ice crystals in which the particles are constantly moving, spinning, and even changing their shapes and sizes due to oscillations of the droplet surface, evaporation, condensation, sublimation, and melting. Although such a particle collection can be treated at each given moment as a fixed cluster, a typical measurement of light scattering takes a finite amount of time over which the spatial configuration of the component particles and their sizes, orientations, and/or shapes continuously and randomly change. Therefore, the registered signal is in effect an average over a large number of distinct clusters.

When a fixed group of particles is illuminated by a monochromatic, spatially co-

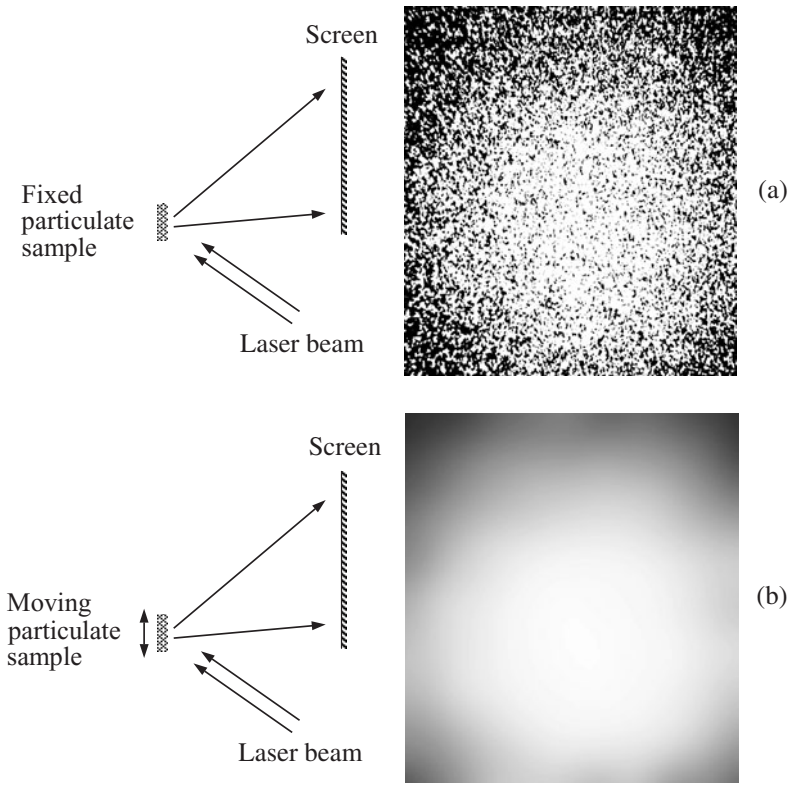


**Figure 1.4.1.** (a) Speckle pattern produced by laser light reflected by a fixed particulate sample. (b) Moving the scattering sample during the measurement averages the speckle pattern out. (After Lenke and Maret, 2000a.)

herent plane wave (e.g., laser light), the light scattered by the group onto a distant screen generates a characteristic speckle pattern consisting of randomly located bright spots of various sizes and shapes (see Fig. 1.4.1(a)). This pattern is the result of constructive or destructive interference of the partial waves scattered by different particles towards a point on the screen. When the particles move, the phase relations between the partial waves constantly change, thereby causing rapid fluctuations of the speckle pattern. Accumulating the signal over a sufficiently long period of time averages the speckle pattern out and results in a rather smooth "incoherent" distribution of the scattered intensity (Fig. 1.4.1(b)).

It has been shown that measurements of the temporal and/or spatial fluctuations of the speckle pattern contain useful information about the particles, in particular about their motion. Statistical analyses of light scattered by dilute and dense particle suspensions, respectively, are the subject of the disciplines called photon correlation spectroscopy (PCS) and diffusing wave spectroscopy (DWS) and form the basis of many well-established experimental techniques for the measurement of various particle characteristics such as velocity, size, and dispersity (e.g., Berne and Pecora, 1976;

Pine *et al.*, 1990). The recent extension of PCS to account for particles changing the polarization state of the incident coherent beam, so-called polarization fluctuation spectroscopy, enables the shapes in addition to the sizes of particles to be sensed (Hopcraft *et al*., 2004).

Photon correlation spectroscopy and diffusing wave spectroscopy study *dynamic* aspects of light scattering by groups of randomly moving particles and as such will not be discussed in this volume. Instead, we will assume that the effect of temporal fluctuations is eliminated by averaging the speckle pattern over a period of time much longer than the typical period of the fluctuations. In other words, we will deal with the average, static component of the scattering pattern. Therefore, the subject of this book can be called *static* light scattering.

## 1.5 Ergodicity

Quantitative analyses of static scattering measurements require the use of a theoretical averaging procedure. Let us consider, for example, the measurement of a scattering characteristic $A$ of a cloud of spherical water droplets. This characteristic depends on time implicitly by being a function of time-dependent physical parameters of the cloud such as the coordinates and sizes of all the constituent particles. The full set of particle positions and sizes will be denoted collectively by $\psi$ and determines the state of the entire cloud at a moment in time. In order to interpret the measurement of $A[\psi(t)]$ accumulated over a period of time extending from $t = t_0$ to $t = t_0 + T$, one needs a way of predicting theoretically the average value

$$\overline{A} = \frac{1}{T}\int_{t_0}^{t_0+T} \mathrm{d}t\, A[\psi(t)]. \tag{1.5.1}$$

Quite often the temporal evolution of a complex scattering object such as the cloud of water droplets is controlled by several physical processes and is described by an intricate system of equations. To incorporate the solution of this system of equations for each moment of time into the theoretical averaging procedure (1.5.1) can be a formidable task and is rarely, if ever, done. Instead, averaging over time is replaced by ensemble averaging based on the following rationale.

Although the coordinates and sizes of water droplets in the cloud change with time in a specific way, the range of instantaneous states of the cloud captured by the detector during the measurement becomes representative of that captured over an infinite period of time provided that $T$ is sufficiently large. We thus have

$$\overline{A} \approx \lim_{\tau\to\infty} \frac{1}{\tau}\int_{t_0}^{t_0+\tau} \mathrm{d}t\, A[\psi(t)] = \langle A\rangle_t. \tag{1.5.2}$$

Notice now that the infinite integral in Eq. (1.5.2) can be expected to "sample" every physically realizable state $\psi$ of the cloud. Furthermore, this sampling is statistically

representative in that the number of times each state is sampled is large and tends to infinity in the limit $\tau \to \infty$. Most importantly, the cumulative contribution of a cloud state $\psi$ to $\langle A \rangle_t$ is independent of the specific moments of time when this state actually occurred in the process of the temporal evolution of the cloud. Rather, it depends on how many times this state was sampled. Therefore, this cumulative contribution can be thought of as being proportional to the probability of occurrence of the state $\psi$ at *any* moment of time. This means that instead of specifying the state of the cloud at each moment $t$ and integrating over all $t$, one can introduce an appropriate time-independent probability density function $p(\psi)$ and integrate over the entire physically realizable range of cloud states:

$$\langle A \rangle_t \approx \int d\psi \, p(\psi) A(\psi) = \langle A \rangle_\psi, \tag{1.5.3}$$

where

$$\int d\psi \, p(\psi) = 1. \tag{1.5.4}$$

The assumption that averaging over time for a "sufficiently random" object can be replaced by ensemble averaging is called the ergodic hypothesis. Although it has not been possible to establish mathematically the full ergodicity of real dynamical systems, more restricted versions of the ergodic theorem have been proven. Physical processes such as Brownian motion and turbulence often help to establish a significant degree of randomness of particle positions and orientations, which seems to explain why many theoretical predictions based on the ergodic hypothesis have agreed very well with experimental data (e.g., Berne and Pecora, 1976). Therefore, we will assume throughout this book that the scattering system in question is ergodic and, thus, Eq. (1.5.3) is applicable.

## 1.6 Single scattering by random particles

The simplest stochastic scattering object is a single particle undergoing random changes of position, orientation, size, and/or shape during the measurement. A good example is a solid or liquid particle trapped inside an electrostatic or optical levitator. In this case particle positions are confined to a small volume with diameter often much smaller than the distance from the volume center to the detector (Fig. 1.6.1). It is then rather straightforward to show that the detector signal accumulated over a period of time is independent of particle positions and can be described in terms of phase and extinction matrices averaged over appropriate ranges of particle orientations, sizes, and shapes. The formalism remains largely the same as in the case of far-field scattering by a fixed object.

A more difficult case is the scattering by a small random group of particles (Fig.