A MICROPHYSICAL MODEL OF SCATTERING, ABSORPTION, AND EXTINCTION IN ELECTROMAGNETIC THEORY

by

MATTHEW JAMES BERG

B.Sc., Colorado School of Mines, Golden, Colorado USA, 2003

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2008

Abstract

This work presents a microphysical model of the classical interaction of electromagnetic waves with arbitrary single and multiple particles. The model is based on the volume integral equation solution to the macroscopic time-harmonic Maxwell equations. The integral is discretized over a particle's volume. The near and far-field scattered wave is then described by the secondary radiation from the discretized elements. The physical origin of the angular structure of the scattered wave is characterized by the superposition of these secondary waves. A graphical technique is developed to visualize how this superposition relates to the physical features of a particle, e.g., its size, shape, and refractive index. Numerical and analytical implementations of the model are presented for spherical and spheroidal particles and fractal-like spherical-particle aggregates. The connection between the reflection symmetry of a particle and the polarization state of its far-field scattered wave is illustrated. The model is used to explain the cause of the angular power-law patterns in a particle's scattered intensity. An analysis of the internal field distribution in fractal-like aggregates is performed and the results are compared to the Rayleigh-Debye-Gans theory. Extinction and the optical theorem are examined within the context of the model, resulting in a new understanding for the physical mechanism causing extinction and implications regarding its measurement. The culmination of this work is the unification of multiple scattering-concepts, often regarded as distinct, and the resulting insight afforded by the unified microphysical picture. This unified view is shown to reveal a new and simple explanation for the famous extinction paradox.

A MICROPHYSICAL MODEL OF SCATTERING, ABSORPTION, AND EXTINCTION IN ELECTROMAGNETIC THEORY

by

MATTHEW JAMES BERG

B.Sc., Colorado School of Mines, Golden, Colorado, USA, 2003

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Physics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2008

Approved by:

Major Professor
Christopher M. Sorensen, Ph.D.

Copyright

Abstract

This work presents a microphysical model of the classical interaction of electromagnetic waves with arbitrary single and multiple particles. The model is based on the volume integral equation solution to the macroscopic time-harmonic Maxwell equations. The integral is discretized over a particle's volume. The near and far-field scattered wave is then described by the secondary radiation from the discretized elements. The physical origin of the angular structure of the scattered wave is characterized by the superposition of these secondary waves. A graphical technique is developed to visualize how this superposition relates to the physical features of a particle, e.g., its size, shape, and refractive index. Numerical and analytical implementations of the model are presented for spherical and spheroidal particles and fractal-like spherical-particle aggregates. The connection between the reflection symmetry of a particle and the polarization state of its far-field scattered wave is illustrated. The model is used to explain the cause of the angular power-law patterns in a particle's scattered intensity. An analysis of the internal field distribution in fractal-like aggregates is performed and the results are compared to the Rayleigh-Debye-Gans theory. Extinction and the optical theorem are examined within the context of the model, resulting in a new understanding for the physical mechanism causing extinction and implications regarding its measurement. The culmination of this work is the unification of multiple scattering-concepts, often regarded as distinct, and the resulting insight afforded by the unified microphysical picture. This unified view is shown to reveal a new and simple explanation for the famous extinction paradox.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank the follow for making this work possible:

- My advisors and teachers at the Colorado School of Mines and Kansas State University who enabled and helped to inspire my ambitions in science. In particular, I want to acknowledge Dr. Chris Sorensen, whose saint-like patience and deeply gifted teaching ability guided much of this work.

- The kind encouragement and critical technical advice given by Dr. Michael Mishchenko. The course and form of this research has been strongly influenced by discussions with Michael and by his own excellent work in electromagnetic theory. I am also grateful for discussions with Drs. Gorden Videen, Thomas Wriedt, Sui-Chun Lee, Anatoli Borovoi, Dan Mackowski, Maxim Yurkin, Kevin Wilson, and Alex Yuffa.

- Adam H. Jundt for crucial assistance in developing a parallelized version of the DDA code used in Ch. 8. Flint Pierce for providing simulated fractal-like aggregates, and Scott Strong for excellent instruction in the C++ programming language.

- Three years of generous support from the NASA Graduate Student Researchers Program, and 100,000 hours of simulation time on the NASA ASD SGI-Altix "Columbia" supercomputer.

Dedication

To my friends and family.

Preface

The phenomena associated with electromagnetic scattering are everywhere. Familiar examples in nature include the blue sky and red sunset, twinkling of stars, rainbows, and glory. In general, electromagnetic scattering refers to the interaction of light with an object; it is ultimately how one is able to see the world around them. Such scattering is not restricted to visible light, but occurs over the entire electromagnetic spectrum. This incites a veritable host of applications in both scientific and technological contexts, including satellite remote sensing of the Earth and the planets, laboratory and field-based investigation of particulate systems *in situ*, wireless communications, and medical imaging, just to name a few. The purpose of this dissertation is to study electromagnetic scattering from an uncommon microphysical perspective. In doing so, a unique, and often graphically-based, description of many familiar scattering effects is achieved; in some cases, this results in a new and profound understanding for the physical causes of these effects.

The classical theory of the propagation and interaction of light with a material body, or particle, is given by the Maxwell equations supplemented with appropriate constitutive relations. In applying these equations to a specific problem, one can choose two distinctly different approaches; macroscopic and microscopic. In the macroscopic approach, the particle is represented by a continuous electromagnetic medium described by a refractive index. This is justified so long as the wavelength of the light is too great to resolve the spatial atomic structure of the material. In the microscopic approach, the electromagnetic medium is regarded as an array of polarizable points in vacuum filling the particle volume. An external wave induces polarization in these points, which then radiate secondary waves that combine with each other and the external wave. The points are thus coupled together, resulting in an infinitely complicated sequence of interactions between them.

Despite their fundamental equivalence, the macroscopic and microscopic approaches do not always yield the same physical insight. The work in this dissertation presents and applies

a model of the classical interaction of light with a particle based on the microscopic approach. From this work, one will see that the microphysical model enables a deeper understanding of the nature of various scattering phenomena. For example, destructive interference within the particle will be seen to be the cause of the angular power-law patterns in a particle's far-field scattered intensity. A new view of extinction is developed, giving the optical theorem a clearer physical meaning, which eventually leads to a new and profoundly fundamental explanation for the extinction paradox. Throughout this work, emphasis is placed on how the behavior of measurable quantities, e.g. Stokes parameters, are related to the field within a particle. Graphical, analytical, and numerical techniques are used to illustrate this connection.

Because of the vast scope of electromagnetic scattering theory and its wide range of applications, it is necessary to clearly define the basic limitations and assumptions that apply throughout the following. These limitations, naturally, follow those that are found in the primary references on which this work builds [1, p. 3-4]:

- The particle is fixed with regard to its location and orientation. This assumption will be valid provided that any motions of an actual particle occur over time scales much shorter than that associated with a measurement.

- All particles reside in vacuum. This is justified since the electromagnetic properties of air in the optical frequency range are far closer to that of vacuum than of typical particle materials.

- Magnetic materials are not considered. This restriction is made partly for simplicity, since it avoids the effects of hysteresis, and partly because the electric response of many materials is dominate over the magnetic response.

- All nonlinear electromagnetic effects are excluded. This will be acceptable provided that fields are weak enough that the electromagnetic properties of a particle are independent of the field strength.

- Particles will be assumed to be uniform, homogeneous, and isotropic. This is done to simplify the mathematical analysis and is not intended to be an implicit restriction in the development or application of the model.

- Only monochromatic electromagnetic waves are considered. This restricts direct application of this work to situations involving ultra-short laser pulse illumination, transient electromagnetic effects, and spatial or temporal incoherence.

- Only elastic scattering is considered, and hence frequency redistribution of the waves is excluded, e.g. Raman scattering and fluorescence.

- Thermal emission from the particle is neglected.

To avoid unnecessary redundancy, and to save space, the reader will often be referred to [1] and [2], the latter of which is freely available on-line in the `.pdf` format. Whenever possible, the same notation and terminology will be used as in these references. The reader should also note that this work is restricted to *classical* electromagnetic theory; no account is made for the quantum mechanical nature of the electromagnetic field, or of the quantum mechanical and statistical aspects of its interaction with matter. Consequently, this work does not constitute a true atomic or molecular model of electromagnetic scattering. The interested reader can consult [3] for the basics of such a treatment of scattering.

Matthew J. Berg

Manhattan, KS

November 2008

Chapter 1

Basic Elements of Electromagnetic Theory

> *"We wish to raise our feeble voice against innovations that can have no other effect than to check the progress of science and renew all those wild phantoms of the imagination which Bacon and Newton put to flight from her temple. We wish to recall philosophers to the strict and severe methods of investigation."*
>
> -H. P. Brougham in [4, p. 29]

This chapter presents aspects of classical electromagnetic theory that are essential for later considerations. This theoretical introduction will establish the foundation on which the microphysical model will be formulated in Chs. 2 and 4, and will present important concepts like energy flow and the Stokes parameters. Many of the limitations and assumptions involved in the microphysical model will be described here, including, e.g., the harmonic time dependence of the fields and sources. A discussion of the definition and meaning of scattering will be given, which will be important for the interpretation of certain scattering effects presented in later chapters.

1.1 Maxwell Equations

The governing differential equations of Maxwell's electromagnetic theory are formulated in terms of a set of four vector fields; the electric field $\boldsymbol{\mathcal{E}}$, magnetic field $\boldsymbol{\mathcal{B}}$, electric displacement

$\boldsymbol{\mathcal{D}}$, and magnetic intensity $\boldsymbol{\mathcal{H}}$. The source of the these fields are electric charge and current density ϱ and $\boldsymbol{\mathcal{J}}$. At a location $\mathbf{r}$ at a time t, the equations are [5, 6]

$$\boldsymbol{\nabla} \times \boldsymbol{\mathcal{E}}(\mathbf{r},t) + \partial_t \boldsymbol{\mathcal{B}}(\mathbf{r},t) = 0, \tag{1.1}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\mathcal{H}}(\mathbf{r},t) - \partial_t \boldsymbol{\mathcal{D}}(\mathbf{r},t) = \boldsymbol{\mathcal{J}}(\mathbf{r},t), \tag{1.2}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{D}}(\mathbf{r},t) = \varrho(\mathbf{r},t), \tag{1.3}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{B}}(\mathbf{r},t) = 0. \tag{1.4}$$

A "particle" in the context of this work is meant to represent any extended material body; it can be arbitrary in size, shape, orientation, and composition so long as it constitutes a linear, uniform, and isotropic electromagnetic medium. The particle's surface is denoted by $\mathcal{S}$ and its internal volume and external region by V^{int} and V^{ext}, respectively. The vector $\mathbf{r}$ denotes the observation point, where the fields values are desired, either inside or outside of the particle. Either Cartesian coordinates (x,y,z), or spherical-polar coordinates (r,θ,ϕ), will be used to describe the observation point. To clarify this notation, Fig. (1.1) schematically shows a particle centered at the coordinate origin along with Cartesian axes x, y, z, and spherical-polar unit vectors $\hat{\boldsymbol{r}}$, $\hat{\boldsymbol{\theta}}$, and $\hat{\boldsymbol{\phi}}$.

Taking the divergence of Eq. (1.2) and combining the result with Eq. (1.3) yields the continuity relation

$$\boldsymbol{\nabla} \cdot \boldsymbol{\mathcal{J}}(\mathbf{r},t) = -\partial_t \varrho(\mathbf{r},t), \tag{1.5}$$

which establishes a connection between the two source distributions and accounts for charge conservation. In vacuum, the fields $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{D}}$, and $\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{H}}$ are related as

$$\boldsymbol{\mathcal{D}}(\mathbf{r},t) = \varepsilon_o \boldsymbol{\mathcal{E}}(\mathbf{r},t), \qquad \boldsymbol{\mathcal{H}}(\mathbf{r},t) = \frac{1}{\mu_o}\boldsymbol{\mathcal{B}}(\mathbf{r},t), \tag{1.6}$$

where ε_o and μ_o are, respectively, the vacuum permittivity and permeability constants, related to the the speed of light $c = (\varepsilon_o \mu_o)^{-1/2}$.

If a particle is exposed to an electromagnetic wave, the fields of the wave will apply forces to the charges in the material. The force on a differential volume element $\mathrm{d}V$ containing

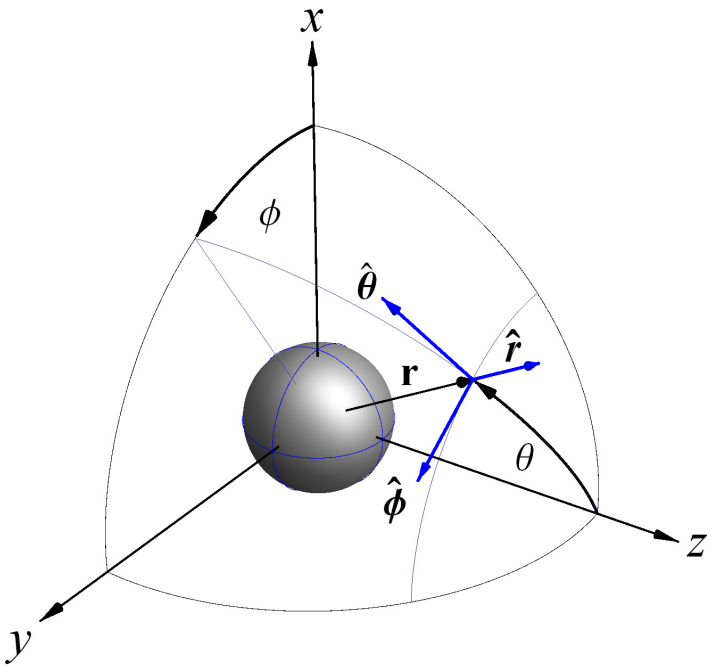

Figure 1.1: Coordinate system with a spherical particle centered at the origin.

charge density ϱ and current density $\boldsymbol{\mathcal{J}}$ is given by the Lorentz force

$$\mathrm{d}\mathbf{F}(\mathbf{r},t) = \left[\varrho(\mathbf{r},t)\boldsymbol{\mathcal{E}}(\mathbf{r},t) + \boldsymbol{\mathcal{J}}(\mathbf{r},t) \times \mathcal{B}(\mathbf{r},\mathbf{t})\right] \mathrm{d}V. \tag{1.7}$$

Depending on the specific properties of the material, this force can result in electric and magnetic polarization. These effects are described by the polarization $\boldsymbol{\mathcal{P}}$ and magnetization $\boldsymbol{\mathcal{M}}$ fields

$$\boldsymbol{\mathcal{P}}(\mathbf{r},t) = \boldsymbol{\mathcal{D}}(\mathbf{r},t) - \varepsilon_o \boldsymbol{\mathcal{E}}(\mathbf{r},t), \tag{1.8}$$

$$\boldsymbol{\mathcal{M}}(\mathbf{r},t) = \frac{1}{\mu_o}\boldsymbol{\mathcal{B}}(\mathbf{r},t) - \boldsymbol{\mathcal{H}}(\mathbf{r},t), \tag{1.9}$$

which represent, respectively, the macroscopic (i.e., averaged) electric and magnetic dipole moment per unit volume induced in the particle material. Both $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{M}}$ are zero in vacuum. If the material is partially conductive, a current density can also be induced via Ohm's law[1]

$$\boldsymbol{\mathcal{J}}(\mathbf{r},t) = \sigma(\mathbf{r},\omega)\boldsymbol{\mathcal{E}}(\mathbf{r},t). \tag{1.10}$$

[1]It will be assumed that Ohm's law applies to all particle materials in this work. Consequently, non-Ohmic materials, such as semiconductors, are not considered here.

Equation (1.10) is one of several constitutive relations that describe how a particle's material responds to the fields in its presence. In this sense, the constitutive relations are supplemental to the Maxwell equations and are required in order to ensure that the Maxwell equations yield a unique and physically meaningful solution. The additional constitutive relations are [6]

$$\boldsymbol{\mathcal{P}}(\mathbf{r},t) = \varepsilon_o \chi_e(\mathbf{r},\omega)\boldsymbol{\mathcal{E}}(\mathbf{r},t), \tag{1.11}$$

$$\boldsymbol{\mathcal{M}}(\mathbf{r},t) = \chi_m(\mathbf{r},\omega)\boldsymbol{\mathcal{H}}(\mathbf{r},t), \tag{1.12}$$

$$\boldsymbol{\mathcal{B}}(\mathbf{r},t) = \mu(\mathbf{r},\omega)\boldsymbol{\mathcal{H}}(\mathbf{r},t), \tag{1.13}$$

$$\boldsymbol{\mathcal{D}}(\mathbf{r},t) = \varepsilon(\mathbf{r},\omega)\boldsymbol{\mathcal{E}}(\mathbf{r},t). \tag{1.14}$$

In Eqs. (1.10)-(1.14), σ is the conductivity, χ_e and χ_m are the electric and magnetic susceptibility, μ is the magnetic permeability, and ε is the electric permittivity of the particle material. Assuming that the material is linear, these quantities are independent of the fields. Moreover, they are constant in time and intrinsic to the material, but are, however, frequency dependent. The electric permittivity and susceptibility are related by

$$\varepsilon(\mathbf{r},\omega) = \varepsilon_o \left[1 + \chi_e(\mathbf{r},\omega)\right] \tag{1.15}$$

which can be seen from the combination of Eqs. (1.8) and (1.11).

1.2 Harmonic Time Dependence

The Maxwell equations (1.1)-(1.4) can be simplified if the time dependence of the fields and their corresponding sources is specified. Further simplification can be made by introducing the concept of Fourier transformed time-harmonic fields. To do this recall that a bounded function $f(t)$ in the time domain can be related to its Fourier transform counterpart $F(\omega)$ in the frequency domain via the integral relations

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) \exp(-i\omega t)\, \mathrm{d}\omega, \tag{1.16}$$

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) \exp(i\omega t)\, \mathrm{d}t. \tag{1.17}$$

Now suppose that the fields and sources in Eqs. (1.1)-(1.4) are described in terms of their frequency domain counterpart. For example [6, Sec. 4.2]

$$\boldsymbol{\mathcal{J}}(\mathbf{r}, t) = \frac{1}{\sqrt{2\pi}} \mathrm{Re} \int_{-\infty}^{\infty} \mathbf{J}(\mathbf{r}, \omega) \exp(-i\omega t)\, \mathrm{d}\omega, \tag{1.18}$$

where $\mathrm{Re}\{\ldots\}$ is the real filter. If the time dependence of the current source $\boldsymbol{\mathcal{J}}$ is harmonic, i.e., has a single angular frequency ω, then the frequency domain current density $\mathbf{J}$ contains only one spectral component. This can be described with the Dirac delta function,

$$\mathbf{J}(\mathbf{r}, \omega) = \mathbf{J}(\mathbf{r}) \sqrt{2\pi} \delta(\omega - \omega'). \tag{1.19}$$

Substituting Eq. (1.19) into Eq. (1.18) gives

$$\boldsymbol{\mathcal{J}}(\mathbf{r}, t) = \int_{-\infty}^{\infty} \mathbf{J}(\mathbf{r}) \delta(\omega - \omega') \exp(-i\omega' t)\, \mathrm{d}\omega'. \tag{1.20}$$

which evaluates to

$$\boldsymbol{\mathcal{J}}(\mathbf{r}, t) = \mathrm{Re}\big[\mathbf{J}(\mathbf{r}) \exp(-i\omega t)\big]. \tag{1.21}$$

In Eq. (1.21), $\boldsymbol{\mathcal{J}}$ is the real-valued (physical) current density, whereas $\mathbf{J}$ describes this same current density, but can now be complex valued. By analogy, the remaining fields and source in the Maxwell equations $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{D}}$, $\boldsymbol{\mathcal{H}}$, and ϱ can be represented in terms of the real part of their corresponding complex-valued time-harmonic frequency-domain quantities, $\mathbf{E}$, $\mathbf{D}$, $\mathbf{B}$, $\mathbf{H}$, and ρ.

There are several advantages to proceeding within the framework of time-harmonic sources and fields. Many laboratory sources are effectively time-harmonic; the bandwidth of the emission from continuous wave laser sources is narrow enough that it can usually be considered monochromatic [7, Ch. 3]. Additionally, nearly any arbitrary time dependence of a source or field distribution can be decomposed into its frequency domain spectrum via Eq. (1.17). The linearity of the Maxwell equations then allows one to solve the equations for each spectral component in the frequency domain and render the full solution in the time domain via superposition and the inverse Fourier relation of Eq. (1.16).

Having now specified the harmonic time dependence of the sources and fields, the time derivatives of the quantities appearing in Eqs. (1.1)-(1.4) can be evaluated, giving

$$\boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) - i\omega \mathbf{B}(\mathbf{r}) = 0, \tag{1.22}$$

$$\boldsymbol{\nabla} \times \mathbf{B}(\mathbf{r}) + i\omega \varepsilon_o \mu_o \mathbf{E}(\mathbf{r}) = \mu_o \big[\mathbf{J}(\mathbf{r}) + \boldsymbol{\nabla} \times \mathbf{M}(\mathbf{r}) - i\omega \mathbf{P}(\mathbf{r}) \big], \tag{1.23}$$

$$\boldsymbol{\nabla} \cdot \mathbf{E}(\mathbf{r}) = \frac{1}{\varepsilon_o} \big[\rho(\mathbf{r}) - \boldsymbol{\nabla} \cdot \mathbf{P}(\mathbf{r}) \big], \tag{1.24}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B}(\mathbf{r}) = 0. \tag{1.25}$$

and the corresponding continuity relation

$$\boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{r}) = i\omega \rho(\mathbf{r}). \tag{1.26}$$

Equations (1.22)-(1.25) are the time-harmonic frequency-domain Maxwell equations. They can be simplified further by generalizing the constitutive parameters σ, χ_e, χ_m, μ, and ε. This generalization consists of redefining these quantities in a piecewise sense to apply both inside and outside of the particle. Then, let

$$\sigma(\mathbf{r}, \omega) = \begin{cases} \sigma(\omega) & \mathbf{r} \in V^{int} \\ 0 & \mathbf{r} \in V^{ext}, \end{cases} \tag{1.27}$$

and χ_e and χ_m be redefined likewise, whereas

$$\varepsilon(\mathbf{r}, \omega) = \begin{cases} \varepsilon(\omega) & \mathbf{r} \in V^{int} \\ \varepsilon_o & \mathbf{r} \in V^{ext}, \end{cases} \tag{1.28}$$

with μ redefined likewise. Since only nonmagnetic particle materials are considered in this work, $\chi_m = 0$. With the time-harmonic frequency-domain versions of Eqs. (1.8) and (1.9), the redefined constitutive parameters, and the constitutive relations of Eqs. (1.10)-(1.12), Eqs. (1.22)-(1.23) above become

$$\boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) = i\omega \mathbf{B}(\mathbf{r}), \tag{1.29}$$

$$\boldsymbol{\nabla} \times \mathbf{B}(\mathbf{r}) = -i\omega \mu_o \epsilon(\mathbf{r}, \omega) \mathbf{E}(\mathbf{r}), \tag{1.30}$$

where

$$\epsilon(\mathbf{r},\omega) = \varepsilon(\mathbf{r},\omega) + \frac{i\sigma(\mathbf{r},\omega)}{\omega} \tag{1.31}$$

is the particle's complex-valued electric permittivity. In deriving Eq. (1.30), Eq. (1.15) is used. One should note the typographical distinction between ε, given by Eq. (1.15), and ϵ given above. The divergence equations (1.24) and (1.25) need not be reconsidered here since each of them can be derived from Eqs. (1.29) and (1.30) by taking the divergence of Eqs. (1.29) and (1.30) and using Eqs. (1.15) and (1.31)[2].

Measurements of the electromagnetic waves typically involve devices that quantify the force on charges in the detector due to the electric field to which it is exposed. Often the frequency of the field is great enough that the detector responds to the time-averaged force exerted by the field rather than the instantaneous force [3]. The time average of a quantity $\mathbf{Q}$ is [2, p. 11]

$$\langle \mathbf{Q}(\mathbf{r}) \rangle_t = \frac{1}{\tau} \int_t^{t+\tau} \mathbf{Q}(\mathbf{r},t')\, \mathrm{d}t', \tag{1.32}$$

where τ is a period much greater than that the period associated a change in $\mathbf{Q}$.

1.3 Wave Equation

The considerations in Secs. 1.1 and 1.2 have reduced the four Maxwell equations of Eqs. (1.1)-(1.4) involving the four field vectors $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{D}}$, $\boldsymbol{\mathcal{B}}$, and $\boldsymbol{\mathcal{H}}$ to the equivalent time-harmonic equations Eqs. (1.29) and (1.30) involving only the fields $\mathbf{E}$ and $\mathbf{B}$. These two equations can be combined into a wave equation following [2, p. 32-33]. For points outside and inside of the particle

$$\left.\begin{aligned} \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) &= i\omega \mathbf{B}(\mathbf{r}) \\ \boldsymbol{\nabla} \times \mathbf{B}(\mathbf{r}) &= -i\omega\mu_o\varepsilon_o \mathbf{E}(\mathbf{r}) \end{aligned}\right\} \mathbf{r} \in V^{ext}, \tag{1.33}$$

$$\left.\begin{aligned} \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) &= i\omega \mathbf{B}(\mathbf{r}) \\ \boldsymbol{\nabla} \times \mathbf{B}(\mathbf{r}) &= -i\omega\mu_o\epsilon \mathbf{E}(\mathbf{r}) \end{aligned}\right\} \mathbf{r} \in V^{int}, \tag{1.34}$$

[2]Note that this does not mean that the Maxwell divergence equations are redundant just because they can be derived from the curl, see [6].

[3]There are certain detection techniques used in ultra-short laser-pulse synthesis that can effectively resolve the full temporal character of and electromagnetic wave at optical frequencies. For example, streak cameras and the polarization gating method in ultra-fast laser pulse optics.

where the frequency dependence of ϵ is suppressed for brevity. Notice that one need only consider the electric field equations in Eqs. (1.33) and (1.34) since the magnetic field can be calculated from the curl of the electric field. Equations (1.33) and (1.34) can be uncoupled by taking the curl,

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) - k^2 \mathbf{E}(\mathbf{r}) = 0 \qquad \mathbf{r} \in V^{ext}, \tag{1.35}$$

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) - m^2 k^2 \mathbf{E}(\mathbf{r}) = 0 \qquad \mathbf{r} \in V^{int}, \tag{1.36}$$

where $k = \omega\sqrt{\mu_o \varepsilon_o}$ and $m = \sqrt{\epsilon/\varepsilon_o}$ are the vacuum wavenumber and complex-valued refractive index, respectively. In addition, $k = 2\pi/\lambda$ where λ is the vacuum wavelength. One can now combine Eqs. (1.35) and (1.36) by introducing the piecewise refractive index

$$\tilde{m}(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in V^{ext} \\ m & \mathbf{r} \in V^{int} \end{cases} \tag{1.37}$$

Then Eq.(1.36) becomes

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) - k^2 \mathbf{E}(\mathbf{r}) = \mathbf{j}(\mathbf{r}) \qquad \mathbf{r} \in V^{ext} \cup V^{int}, \tag{1.38}$$

where

$$\mathbf{j}(\mathbf{r}) = k^2 \left[\tilde{m}^2(\mathbf{r}) - 1\right] \mathbf{E}(\mathbf{r}), \tag{1.39}$$

is called a forcing function, which acts like a source term for the wave equation.

Equation (1.38) is the time-harmonic frequency-domain Maxwell wave equation and contains the same information as the four equations (1.1)-(1.4). The particle's material is assumed to constitute a linear, isotropic, and homogeneous macroscopic electromagnetic medium making m a constant inside the particle and unity outside. Note that m essentially quantifies the response of the particle material to the spectral component ω, hence the solutions to Eq. (1.38) will in general depend on frequency.

Notice that the fields $\mathbf{D}$ and $\mathbf{H}$ have been eliminated. These fields are replaced by the introduction of the polarization $\mathbf{P}$ and magnetization $\mathbf{M}$ fields and are related to $\mathbf{E}$ and $\mathbf{B}$ through the constitutive relations of Eqs. (1.11)-(1.13). However, since $\mathbf{B}$ is proportional to the curl of $\mathbf{E}$, the electromagnetic state of a particle can be completely described in terms

of electric field residing in its interior. Consequently, the solutions of the wave equation (1.38) developed later will focus on expressions for the electric field only. An additional advantage of eliminating $\mathbf{D}$ and $\mathbf{H}$ in favor of $\mathbf{E}$ and $\mathbf{B}$ is that it is these latter fields that appear in the Lorentz force of Eq. (1.7). This is appealing because it can provide one with a mechanical-like perspective for the origin of the response of a particle to an external field [6, Sec. 2.4].

Notice that the terms "electromagnetic field" and "electromagnetic wave," are used essentially interchangeably. In a technical sense, the electromagnetic field is the mathematical object consisting of the four vector fields $\boldsymbol{\mathcal{E}}$, $\boldsymbol{\mathcal{B}}$, $\boldsymbol{\mathcal{D}}$, and $\boldsymbol{\mathcal{H}}$, or equivalently $\mathbf{E}$, $\mathbf{B}$, $\mathbf{D}$, and $\mathbf{H}$, which obey the Maxwell equations (1.1)-(1.4) or (1.22)-(1.25), respectively [5]. As such, the electromagnetic field is a general term that describes all the fields given by the Maxwell equations for an *arbitrary* source distribution. Then, what is meant by the term "electromagnetic wave" is the collection of fields obeying the Maxwell equations for sources that result in wave-like solutions, i.e. Eq. (1.38). These wave-like solutions are often transverse and can therefore be described by the electric field component of the wave only. Consequently, "wave" and "field" are often interchangeable terms; "wave" will be used in the following when it is important to emphasize the role of both the electric and magnetic fields, whereas "field" will be used otherwise.

1.4 Boundary Conditions

A subtle property of the Maxwell equations is that they are valid only in regions where the constitutive parameters are continuous [6, Sec. 2.8]. Consequently, the Maxwell equations will yield different solutions for $\mathbf{E}$ and $\mathbf{B}$ inside and outside the particle. The fields in these regions must then be connected. The electromagnetic boundary conditions provide this connection between the fields on either side of the particle's surface, and hence connect the fields in the two regions V^{int} and V^{ext}. Expressions for the boundary conditions are obtained by casting the Maxwell equations in integral form and examining the limiting behavior of

the fields over closed surfaces that straddle the discontinuity at $\mathcal{S}$ [6, Sec. 2.8]. These expressions are given in terms of the components of the field that are tangential and normal to $\mathcal{S}$,

$$\hat{\mathbf{n}} \times \left[\mathbf{B}^{ext}(\mathbf{r}) - \mathbf{B}^{int}(\mathbf{r})\right] = \begin{cases} \mu_o \mathbf{K}(\mathbf{r}) & \text{perfect conductor: } \sigma \to \infty \\ 0 & \text{otherwise} \end{cases} \tag{1.40}$$

$$\hat{\mathbf{n}} \times \left[\mathbf{E}^{ext}(\mathbf{r}) - \mathbf{E}^{int}(\mathbf{r})\right] = 0 \tag{1.41}$$

$$\hat{\mathbf{n}} \cdot \left[\mathbf{B}^{ext}(\mathbf{r}) - \mathbf{B}^{int}(\mathbf{r})\right] = 0 \tag{1.42}$$

$$\hat{\mathbf{n}} \cdot \left[\varepsilon_o \mathbf{E}^{ext}(\mathbf{r}) - \varepsilon \mathbf{E}^{int}(\mathbf{r})\right] = \rho_s(\mathbf{r}), \tag{1.43}$$

where $\mathbf{r} \in \mathcal{S}$ and $\hat{\mathbf{n}}$ is the particle's outward surface-normal, see Fig. (1.2). Equation (1.14) and the assumed nonmagnetic nature of the particle are used in deriving Eqs. (1.40)-(1.43). If a particle is perfectly conducting, i.e. $\sigma \to \infty$, any current density is confined to the surface and hence becomes a surface current density $\mathbf{K}$. This current corresponds to a discontinuity in the magnetic field as indicated in Eq. (1.40). If however, the particle is partially conductive or non-conductive, then there is no surface current and the magnetic field is continuous across the particle surface [2].

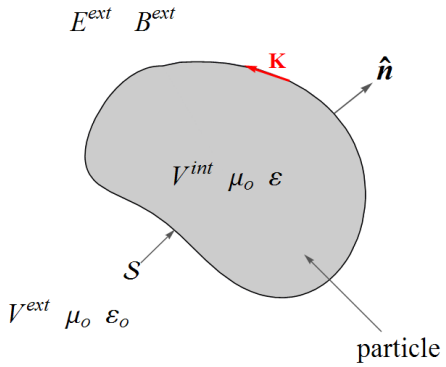

Figure 1.2: Perfectly conducting particle with surface current $\mathbf{K}$.

1.5 Energy

The electromagnetic field contains energy. This energy content is present in the field regardless of whether the field is static or time dependent. If the field propagates, energy can

be transported between regions, constituting an energy flow[4]. This flow is described by the Poynting vector

$$\boldsymbol{\mathcal{S}}(\mathbf{r},t) = \boldsymbol{\mathcal{E}}(\mathbf{r},t) \times \boldsymbol{\mathcal{H}}(\mathbf{r},t), \tag{1.44}$$

which can be derived from the conservation of energy and has dimensions of energy per unit time and area, see [5, Sec. 2.19] and [15, Sec. 6.7].

Note that the Poynting vector does not specify where the energy is contained: Is the energy distributed throughout the field or is it somehow contained within the charged matter composing the sources of the field? This question is discussed in [25, p. 27-4] and is currently an unsolved mystery in theoretical physics. This issue, however, is essentially irrelevant for the considerations of this work. What is important is that the principle of energy conservation apply not only to all the sources and fields of a system *collectively*, but also *locally* within differential volume elements at arbitrary locations. This gives the Poynting vector genuine physical meaning in a measurement context and will be important for the concepts of extinction and scattering later on.

The time-harmonic frequency-domain equivalent of Eq. (1.44) can be obtained by substituting the frequency-domain fields $\mathbf{E}$ and $\mathbf{B}$ into Eq. (1.44) using Eq. (1.13) and exploiting the nonmagnetic nature of the particle, i.e. $\mu = \mu_o$. Then, the time-averaged frequency-domain Poynting vector follows from Eq. (1.32) as [5, p.136-137]

$$\langle \mathbf{S}(\mathbf{r}) \rangle_t = \frac{1}{\mu_o} \mathrm{Re}\left\{ \mathbf{E}(\mathbf{r}) \times [\mathbf{B}(\mathbf{r})]^* \right\}, \tag{1.45}$$

where the asterisk denotes complex conjugation and $\langle \ldots \rangle_t$ represents time averaging. The component of $\langle \mathbf{S}(\mathbf{r}) \rangle_t$ directed *into* the face of an ideal detector is taken to describe the response of that detector to the electromagnetic field, see [1, Sec 3.6].

One should note that the electromagnetic field also carries momentum, see [15, p. 262]. Conservation of this momentum as the field interacts with a particle shows that the particle

[4]There is an enduring controversy regarding exactly how to describe the energy content and flow, e.g. see [5, p. 134], [8, p.25-26], [9–14]. Despite this, the traditional interpretations are pursued here since they are well-verified by measurement, [15, p. 258-262][6, p.85-88][16–23][24, p. 7-10].

experiences forces in the form of radiation pressure and torque. This field momentum will not be considered in this work because its affects on a particle are usually negligible for typical laboratory field strengths.

1.6 Stokes Parameters

Polarization is a fundamental property of *all* electromagnetic waves due to the vector character of its constituent fields. Polarization is usually described by the instantaneous direction of the wave's electric field. In the context of measurement, nearly all real-life detectors are incapable of resolving the temporal changes of a wave at optical frequencies; instantaneous polarization measurement are almost never done. Consequently, a proper description polarization corresponding to the response of a detector should involve only time-averaged power quantities. In the framework of time-harmonic transverse electromagnetic waves, this measurement-oriented description is given by the four Stokes parameters, I, Q, U, and V. Expressions for the Stokes parameters can be formulated for any time-harmonic wave provided that it is *transverse* [2, Sec. 1.3].

Consider a plane wave traveling in vacuum with electric field

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_o \exp(ikr\hat{\mathbf{n}} \cdot \hat{\mathbf{r}}), \tag{1.46}$$

where $\hat{\mathbf{n}}$ is the propagation direction and $\mathbf{E}_o$ is the complex-valued wave amplitude. If the components of the vector amplitude $\mathbf{E}_o$ are given in spherical-polar coordinates, the four Stokes parameters are [2]

$$I(\hat{\mathbf{r}}) = \frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\left(E_{o\theta}E^*_{o\theta} + E_{o\phi}E^*_{o\phi}\right), \tag{1.47}$$

$$Q(\hat{\mathbf{r}}) = \frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\left(E_{o\theta}E^*_{o\theta} - E_{o\phi}E^*_{o\phi}\right), \tag{1.48}$$

$$U(\hat{\mathbf{r}}) = -\frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\left(E_{o\theta}E^*_{o\phi} + E_{o\phi}E^*_{o\theta}\right), \tag{1.49}$$

$$V(\hat{\mathbf{r}}) = \frac{i}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\left(E_{o\phi}E^*_{o\theta} - E_{o\theta}E^*_{o\phi}\right). \tag{1.50}$$

Here $E_{o\theta} = \hat{\boldsymbol{\theta}} \cdot \mathbf{E}_o$ and $E_{o\phi} = \hat{\boldsymbol{\phi}} \cdot \mathbf{E}_o$, where $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ are the spherical-polar unit vectors shown in Fig. (1.1). Notice that the Stokes parameters, defined this way, are independent of the distance from the source of the wave, i.e., they depend only on direction, $\hat{\mathbf{r}}$. This will be convenient later for describing the polarization state of a wave at large distances from a particle. The definition of I, Q, U, and V in Eq. (1.47)-(1.50) is not necessarily restricted to plane waves and will be applied to spherical waves in Sec. 6.1.2.

Each of the Stokes parameters are time-averaged quantities that have dimensions of power per area and can be measured using combinations of linear polarizers and waveplates, see [1, Sec. 2.10]. Knowledge of all four parameters constitutes a full description of the polarization state of a transverse wave. This description can be visualized using using the ellipsometric interpretation of the Stokes parameters, see [1, Sec. 2.7]. In this visualization, an arbitrary polarization state is described by the ellipse, called the vibration ellipse, that is traced out in the $\hat{\boldsymbol{\theta}}$-$\hat{\boldsymbol{\phi}}$ plane at $\mathbf{r}$ by the tip of the wave's electric field vector during one complete cycle of oscillation, see Fig. (1.3). There are four characteristics of the vibration ellipse; its overall size I, its handedness η, ellipticity e, and orientation angle ζ. The handedness indicates the sense of rotation of the electric field and is right-handed if the field rotates in the counter-clockwise direction to an observer looking into the oncoming wave. It is left-handed *vice versa*. If $\eta > 0$ ($\eta < 0$) the electric vector rotates in the right-handed (left-handed) sense. If $\eta = 0$, the wave is linearly polarized and has no handedness. The ellipticity e describes the ratio of the minor and major axes of the vibration ellipse and can be related to the Stokes parameters through the angle[5] β [2, Sec. 2.4]

$$\tan 2\beta(\hat{\mathbf{r}}) = \frac{-V(\hat{\mathbf{r}})}{\sqrt{Q(\hat{\mathbf{r}})^2 + U(\hat{\mathbf{r}})^2}}. \tag{1.51}$$

Then,

$$e(\hat{\mathbf{r}}) = |\tan\beta(\hat{\mathbf{r}})|\,. \tag{1.52}$$

When $e = 0$ the vibration ellipse is a line and the wave is linearly polarized. When $e = 1$

[5]Note that the angle β in this context should not be confused with the Euler rotation angle presented in Sec. 5.4.

the major and minor axes of the ellipse are equal and the wave is circularly polarized. The handedness η is related to β as

$$\eta(\hat{\mathbf{r}}) = \begin{cases} 1 & \beta(\hat{\mathbf{r}}) > 0 \text{ right} \\ 0 & \beta(\hat{\mathbf{r}}) = 0 \text{ linear} \\ -1 & \beta(\hat{\mathbf{r}}) > 0 \text{ left.} \end{cases} \tag{1.53}$$

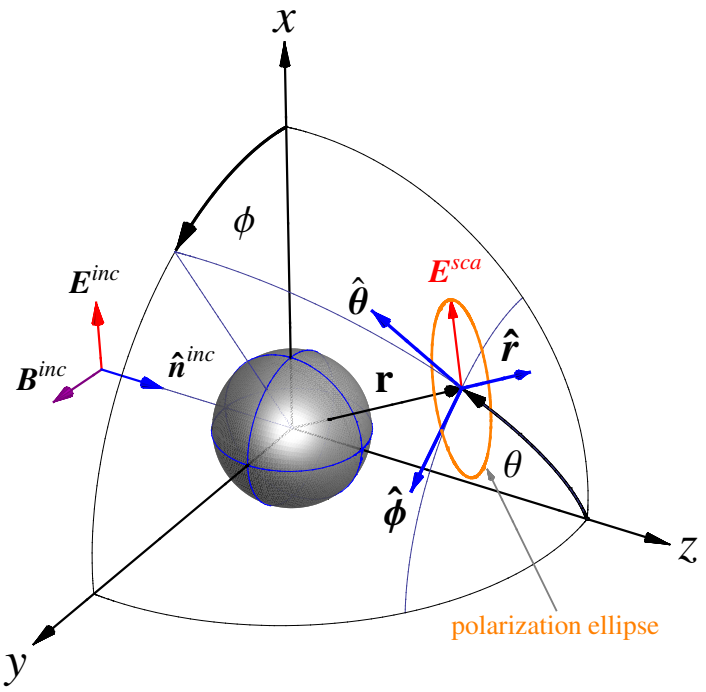

Figure 1.3: The vibration ellipse of the scattered electric field $\mathbf{E}^{sca}$ in the far-field zone, see Sec. 4.2.

In addition to its ellipticity and handedness, the vibration ellipse has an orientation angle ζ with respect to the $\hat{\boldsymbol{\theta}}$ direction at $\mathbf{r}$, see [2, p. 21]. The orientation angle ζ is related to the Stokes parameters as

$$\tan 2\zeta(\hat{\mathbf{r}}) = -\frac{U(\hat{\mathbf{r}})}{Q\hat{\mathbf{r}}}\,. \tag{1.54}$$

In general, the handedness, ellipticity, and orientation of the vibration ellipse of a particle's scattered wave will change with direction.

It is common to measure only the wave intensity I. This is likely due to the substantial additional difficulty of incorporating the required optical elements needed to measure Q, U, and V. While it is true that I contains information regarding the physical character of the

particle under observation, one should not forget that useful information is contained in the other Stokes parameters as well. This fact will be an important motivation for the work presented in Ch. 6.

A related issue concerns the measurement of Stokes parameters for a system of single or multiple particles that undergo random motion during the course of the measurement. In this situation, it is possible that Q, U, and V average to zero, with I being the only nonzero parameter. It is common then to refer to the wave as being "unpolarized," since it appears that it no longer has a directional character. However, this terminology is misleading since *an electromagnetic wave is never unpolarized*; at every instant in time, the electric and magnetic fields have a single and unambiguous direction.

1.7 Comments on the Concept of Scattering

It is important at this point to clearly define the concept of electromagnetic *scattering* as it relates to this work. Different views for the physical meaning of scattering in this context are common despite the fact that they all ultimately reside on the same theoretical foundation; the Maxwell equations. A particularly misleading concept is that of the photon, at least in as much as it is used to describe the *propagation* of light [26]. Here a brief discussion of the meaning of scattering in the Maxwell theory and the role of the photon concept will be presented. Note that the following is based on [27], and the reader interested in greater detail is referred to that reference.

Consider a volume of space centered on the coordinate origin, called the scattering volume, that will contain a particle. Now suppose that the particle is absent and an electromagnetic wave occupies all of space, including the scattering volume. This wave is called the incident wave and the sources that create it will be assumed to be unaffected by the presence of a particle. Such a wave might represent a laser with a beam width that is much greater than the dimensions of the scattering volume. The electric and magnetic fields at each point are solutions of the Maxwell equations (1.1)-(1.4), and the response of a detector

is given by the time-averaged Poynting vector of Eq. (1.44).

If one now introduces a particle in the scattering volume, the wave at the detector will change. The interaction of the fields of the incident wave with the particle establishes a new wave, called the total wave, and it is this total wave that the detector responds to. Because the Maxwell equations are liner, the total wave can be *mathematically* separated into the incident and scattered waves;

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \mathbf{E}^{sca}(\mathbf{r}).$$

Said differently, the scattered wave is the difference between the total wave at $\mathbf{r}$ when the particle is present and when it is not. *This is the meaning of scattering in classical electromagnetic theory* [27]. One must remember that when the particle is present, the incident wave no longer exists and the separation of the total wave into the "incident-plus-scattered" is a mathematical separation that cannot be directly resolved by a detector.

It is possible however, to arrange a measurement so that there is a certain *real* separation between the incident and scattered waves. Such a separation results if the width of the laser beam constituting the incident wave is small enough that the detector can be rotated to locations that are outside of the beam. When the detector is inside the beam, the total wave includes both the incident and scattered waves. When it is outside, the incident wave vanishes and only the scattered wave remains. In this sense, it is essentially possible to measure a particle's scattered wave directly. However, care must be exercised in assuming that this separation can always be done. For example, Ch. 10 shows that the two waves are never separate in the forward and backward directions, and the interference between the two is essential for understanding the meaning of extinction and the related optical theorem.

Occasionally, analogies between electromagnetic scattering and a mechanical collision-based view of scattering are made. Such views, reminiscent of Newton, describe light propagation as a collection of corpuscles, or "photons," that travel in straight line trajectories. Upon encountering a particle, the photons can be absorbed or scattered by the particle. If scattered, the course of the photon trajectory is changed, and thus the collection of all such

photons describes the angular redistribution of light observed in scattering experiments. Taking this view of scattering theory is not only in violation of the Maxwell equations, but can also generate a mindset that, in some cases, can be severely misleading. A classic example of the confusion that can result from the photon view is that of the extinction paradox, see Ch. 11.

The reality is that the *classical* Maxwell theory of light is a *wave* theory, which is clear from Eq. (1.38). Light propagation is described in terms of the concepts of wave motion; the ray-like behavior of the propagation is a limiting behavior of this wave nature as $\lambda \to 0$. The "photon" terminology is misleading in that it tacitly neglects this inherent wave nature. The true meaning of a photon relates to the energy quantization of the electromagnetic field; it is associated with a plane wave component of the frequency-domain Fourier representation of the fields' vector potential [28, Sec. 23], [29]. The photon has a single energy $\hbar\omega$ and propagation direction $\hat{\mathbf{k}}$. Its physical meaning in a strict sense is that it describes the quantized nature in which light *interacts* with a particle; an atom absorbs or emits an photon of energy $\hbar\omega$ and momentum $\mathbf{k}$. The details of this interaction fall within the proper domain of quantum electrodynamic theory and are beyond the scope of this work. With regard to the *propagation* of light, the photon is not to be interpreted as a "particle of light." While it is true that the photon has a single momentum vector $\mathbf{k}$ and behaves like a light ray, its formal representation as a plane wave means that it has an *infinite transverse extent*: The photon hardly constitutes a localized entity.

Chapter 2

Solutions of the Wave Equation

"...there is an infinite amount to convey, it would take me an infinite amount of time..."
-D. M. Wood

Chapter 1 demonstrates that the Maxwell equations can be combined to yield a wave equation for the total electric and magnetic field corresponding to an arbitrary particle. Here, the Green's function and series-based solution methods are presented. An important element in the Green's function method is the Volume Integral Equation (VIE), which will be seen in Ch. 4 to form the analytical foundation for the microphysical model. The VIE is used below to solve the wave equation for the two important examples of spherical and planar waves. Properties of these waves that are important for later considerations are discussed. The reader should note that much of the notation, terminology, and content of this chapter closely follows [1, 2] and is reproduced here for completeness and reference in the remainder of the work.

2.1 Green's Functions Method

Before describing how the wave equation is solved using the Green's function method, it is convenient to introduce the compact and mathematically useful notation of a dyadic. A dyadic can be viewed as an operator that that rotates and scales a vector, see [1, App. A]. The vector is real or complex-valued, and the dyadics take the form of 3×3 matrices. For

example, the identity dyadic is

$$\overleftrightarrow{\mathbf{I}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.1}$$

where each column of the matrix represents a unit vector of the coordinate system[1].

Now recall Eq. (1.38), and replace the source term on the right-hand side by the dyadic unit source $\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')$. The result is the defining equation for the electric dyadic Green's function $\overleftrightarrow{\mathbf{G}}_e$ [30, p. 59],

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') - k^2\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}'). \tag{2.2}$$

This equation can be solved several ways, e.g. see [30, Sec. 4.2] and [15, Sec. 6.4], giving

$$\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \left(\overleftrightarrow{\mathbf{I}} + \frac{1}{k^2}\boldsymbol{\nabla}\otimes\boldsymbol{\nabla}\right) g(\mathbf{r},\mathbf{r}'), \tag{2.3}$$

where

$$g(\mathbf{r},\mathbf{r}') = \frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{2.4}$$

is the scalar retarded Green's function[2] and the notation $\boldsymbol{\nabla}\otimes\boldsymbol{\nabla}$ denotes the dyadic product between the vectors $\boldsymbol{\nabla}$, see App. A.

Now return to the wave equation (1.38) and consider what it becomes if the particle is absent. This situation is realized by replacing the distribution of refractive index, described by $\tilde{m}$ in Eq. (1.37), by unity throughout V^{int}. The result is the homogeneous wave equation and the field entering in it corresponds to the incident field. This incident field is produced by the impressed sources discussed in Sec. 1.7 and is denoted $\mathbf{E}^{inc}$;

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{E}^{inc}(\mathbf{r}) - k^2\mathbf{E}^{inc}(\mathbf{r}) = 0 \qquad \mathbf{r}\in V^{ext}\cup V^{int}. \tag{2.5}$$

An explicit form for this incident field and its corresponding source distribution are given in Sec. 2.3.2.

[1]The vector upon which a dyadic operates must be expressed in the same coordinate system as the dyadic.

[2]Note that there is another solution, the advanced Green's function, which is mathematically equivalent to Eq. (2.4) and consists of the complex conjugate of this equation.

The solution to the wave equation when the particle is present will consist of the superposition of the incident and scattered fields. An expression for $\mathbf{E}^{sca}$ in terms of the electric dyadic Green's function is obtained by operating both sides of Eq. (2.2) on the source function $\mathbf{j}$ of Eq. (1.38),

$$\left(\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{I}}-k^2\overleftrightarrow{\mathbf{I}}\right)\cdot\left[\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{j}(\mathbf{r}')\right]=\overleftrightarrow{\mathbf{I}}\cdot\mathbf{j}(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}'),\tag{2.6}$$

where use has been made of Eq. (A.1). Next, Eq. (2.6) can be integrated over all space to give

$$\left(\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{I}}-k^2\overleftrightarrow{\mathbf{I}}\right)\cdot\int_{\mathbb{R}^3}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{j}(\mathbf{r}')\,\mathrm{d}\mathbf{r}'=\mathbf{j}(\mathbf{r}),\tag{2.7}$$

where $\mathbb{R}^3=V^{ext}\cup V^{int}$ represents all space. Comparing this integral expression to the wave equation

$$\left(\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{I}}-k^2\overleftrightarrow{\mathbf{I}}\right)\cdot\mathbf{E}(\mathbf{r})=\mathbf{j}(\mathbf{r}),$$

shows that

$$\mathbf{E}^{sca}(\mathbf{r})=\int_{V^{int}}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{j}(\mathbf{r}')\,\mathrm{d}\mathbf{r}'.\tag{2.8}$$

Note that the integral in Eq. (2.7) is evaluated over the particle volume V^{int} only since the source function $\mathbf{j}$ vanishes in V^{ext}, see Eq. (1.39). It is tacitly assumed in arriving at Eq. (2.8) by comparison to the wave equation for the *total* field, that the incident field must still be taken into account before a complete solution is obtained. Using Eq. (1.39) one can express Eq. (2.8) as

$$\mathbf{E}^{sca}(\mathbf{r})=k^2(m^2-1)\int_{V^{int}}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}^{int}(\mathbf{r}')\,\mathrm{d}\mathbf{r}'\quad\mathbf{r}\in V^{ext}\cup V^{int},\tag{2.9}$$

where $\mathbf{E}^{int}$ is the particle's internal electric field. The full solution to the wave equation (1.38) for the total electric field follows from the superposition of the incident and scattered fields

$$\mathbf{E}(\mathbf{r})=\mathbf{E}^{inc}(\mathbf{r})+k^2(m^2-1)\int_{V^{int}}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}^{int}(\mathbf{r}')\,\mathrm{d}\mathbf{r}'\quad\mathbf{r}\in V^{ext}\cup V^{int}.\tag{2.10}$$

Equation (2.10) is the volume integral equation for the electric field. Notice that the VIE does not solve the wave equation since the internal field is not yet known. One way that a

solution can be found is to assume that the internal field is the same as the incident field, evaluate the VIE, then substitute the result back in the VIE for the internal field. Repeating this process allows one to iteratively solve the VIE and is the basis for an order-of-scattering expansion for the total field [1, p. 71]. Another method involves discretization of the volume integral to form a system of coupled linear algebraic equations. An iterative numerical procedure is then used to solve the system of equations. This forms the mathematical basis for the Discrete Dipole Approximation (DDA) presented in Ch. 5.

The electric dyadic Greens's function can be given a physical meaning. Recall that the equation defining the function [Eq. (2.2)], is equivalent in form to the wave equation (1.38) except that $\mathbf{j}$ is replaced by the unit source $\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')$. This substitution corresponds to replacing the arbitrary distribution of induced sources in the particle by three point sources of unit magnitude located at $\mathbf{r}'$, where each source is directed along a Cartesian axis. The Green's function then represents the solution to the wave equation for the electric field due to the unit point sources, and the nature of this solution will be examined in Sec. 2.3.1. The linearity of the Maxwell equations then permits formulation of the full solution of the wave equation for the arbitrary source $\mathbf{j}$ as a linear superposition of the solutions corresponding to the unit sources. The mathematical embodiment of this superposition is given by Eq. (2.8), eventually leading to (2.10).

Because $\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')$ represents the unit source corresponding to the electric field, the source appropriate for the magnetic field follows from the curl. To see this, one can take the curl of the wave equation (1.38) and use Eq. (1.29) to obtain

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{B}(\mathbf{r})-k^2\mathbf{B}(\mathbf{r})=-\frac{i}{\omega}\boldsymbol{\nabla}\times\mathbf{j}(\mathbf{r}). \tag{2.11}$$

Then, in analogy to Eq. (2.2), the defining equation for the magnetic dyadic Green's function is [30, p. 59]

$$\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}')-k^2\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}')=\boldsymbol{\nabla}\times\left[\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')\right]. \tag{2.12}$$

The solution for $\overleftrightarrow{\mathbf{G}}_m$ follows from the curl of Eq. (2.3)

$$\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') = \boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}'), \tag{2.13}$$

which gives

$$\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') = \left[\boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{I}} + \frac{1}{k^2}\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \otimes \boldsymbol{\nabla})\right] g(\mathbf{r},\mathbf{r}')$$

and can be simplified using Eq. (A.2)-(A.4) to yield

$$\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') = \boldsymbol{\nabla} g(\mathbf{r},\mathbf{r}') \times \overleftrightarrow{\mathbf{I}}. \tag{2.14}$$

Now consider the wave equation for the magnetic field re-expressed as

$$\left(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{I}} - k^2 \overleftrightarrow{\mathbf{I}}\right) \cdot \mathbf{B}(\mathbf{r}) = -\frac{i}{\omega}\boldsymbol{\nabla} \times \mathbf{j}(\mathbf{r}),$$

and compare this to the integral over all space of Eq. (2.12) operating on $\mathbf{j}$,

$$\begin{aligned}\left(\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{I}} - k^2 \overleftrightarrow{\mathbf{I}}\right) \cdot \int_{\mathbb{R}^3} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{j}(\mathbf{r}')\,\mathrm{d}\mathbf{r}' &= \int_{\mathbb{R}^3} \boldsymbol{\nabla} \times \left[\overleftrightarrow{\mathbf{I}}\delta(\mathbf{r}-\mathbf{r}')\right] \cdot \mathbf{j}(\mathbf{r}')\,\mathrm{d}\mathbf{r}', \\ &= \boldsymbol{\nabla} \times \mathbf{j}(\mathbf{r}).\end{aligned} \tag{2.15}$$

This comparison shows that

$$\mathbf{B}^{sca}(\mathbf{r}) = -\frac{i}{\omega}\int_{V^{int}} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{j}(\mathbf{r}')\,\mathrm{d}\mathbf{r}'. \tag{2.16}$$

Equation (2.16) can be expressed in terms of the particle's internal *electric* field as

$$\mathbf{B}^{sca}(\mathbf{r}) = -i\omega\mu_o\varepsilon_o(m^2-1)\int_{V^{int}} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{E}^{int}(\mathbf{r}')\,\mathrm{d}\mathbf{r}', \quad \mathbf{r} \in V^{ext} \cup V^{int}, \tag{2.17}$$

which yields the full solution to the wave equation for the total magnetic field

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}^{inc}(\mathbf{r}) - i\omega\mu_o\varepsilon_o(m^2-1)\int_{V^{int}} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{E}^{int}(\mathbf{r}')\,\mathrm{d}\mathbf{r}', \quad \mathbf{r} \in V^{ext} \cup V^{int}. \tag{2.18}$$

2.2 Series Expansion Method

An alternative to the Green's function integral-based method to solve the wave equation is to expand the solution in a series of complete and orthogonal functions with known mathematical properties. This series-based method is the most commonly used method in practice. One reason for this is because, for certain particle shapes, exact and analytical solutions can be found. A famous example is the Mie solution which applies to uniform spherical particles. More recently the T-Matrix method of [31], which is based on the series solution, has been improved to the point where solutions can be found for uniform, smooth, and convex-shaped particles to arbitrary numerical accuracy, see [2, Part II]. Another reason for the popularity of the series-based method is that the time required to perform these calculations are often substantially less than numerically solving the alternative integral equations.

Although the series method is typically faster than the Green's function method, it is more difficult to gain insight as to the physical origin of the character of the solution. As will be seen below, the orthogonal functions involved in the series-based solution do not necessarily have any connection to the physical nature of the particle, i.e., its size, shape, orientation, etc. It is only after the boundary conditions of Eq. (1.40)-(1.43) are used to find the expansion coefficients in the series does the solution become unique to the specific particle under consideration. However, because of its speed and capability, the series-based solution will be used frequently in this work for verification purposes, and in some cases, to feed the VIEs with the internal fields needed by the microphysical model.

The orthogonal functions used in the series-based solution are the Vector Spherical Wave Functions (VSWFs), denoted Rg$\mathbf{M}$, Rg$\mathbf{N}$, $\mathbf{M}$, and $\mathbf{N}$. These functions form a complete set upon which any solution to the time-harmonic frequency-domain Maxwell wave equation can be expanded. Because the wave equation is a second order partial differential equation, the set consists of *pairs* of orthogonal functions. The function pair Rg$\mathbf{M}$ and Rg$\mathbf{N}$ are finite everywhere, which makes them suited for expansion of the incident and internal fields [2,

Sec. 5.1]

$$\mathbf{E}^{inc}(\mathbf{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[a_{mn}\mathrm{Rg}\mathbf{M}_{mn}(k\mathbf{r}) + b_{mn}\mathrm{Rg}\mathbf{N}_{mn}(k\mathbf{r})\right] \quad \mathbf{r}\in V^{int}\cup V^{ext}, \tag{2.19}$$

$$\mathbf{E}^{int}(\mathbf{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[c_{mn}\mathrm{Rg}\mathbf{M}_{mn}(k^{int}\mathbf{r}) + d_{mn}\mathrm{Rg}\mathbf{N}_{mn}(k^{int}\mathbf{r})\right] \quad \mathbf{r}\in V^{int}. \tag{2.20}$$

The functions $\mathbf{M}$ and $\mathbf{N}$ obey the Sommerfeld radiation condition at $kr\to\infty$, diverge at the origin, but vanish at infinity and hence are suited for expansion of the scattered wave

$$\mathbf{E}^{sca}(\mathbf{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[p_{mn}\mathbf{M}_{mn}(k\mathbf{r}) + q_{mn}\mathbf{N}_{mn}(k\mathbf{r})\right], \quad \mathbf{r}\in V^{ext}. \tag{2.21}$$

Note the appearance of the refracted wavenumber $k^{int} = mk$ in the expansion for the internal wave Eq. (2.20), where m here represents the refractive index and should not be confused with the index m.

Using Eq. (1.30), the VSWF expansion for the incident, internal, and scattered magnetic fields follow from Eqs. (2.19)-(2.21)

$$\mathbf{B}^{inc}(\mathbf{r}) = -\frac{ik}{\omega}\sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[a_{mn}\mathrm{Rg}\mathbf{N}_{mn}(k\mathbf{r}) + b_{mn}\mathrm{Rg}\mathbf{M}_{mn}(k\mathbf{r})\right] \quad \mathbf{r}\in V^{int}\cup V^{ext}, \tag{2.22}$$

$$\mathbf{B}^{int}(\mathbf{r}) = -\frac{ik^{int}}{\omega}\sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[c_{mn}\mathrm{Rg}\mathbf{N}_{mn}(k^{int}\mathbf{r}) + d_{mn}\mathrm{Rg}\mathbf{M}_{mn}(k^{int}\mathbf{r})\right] \quad \mathbf{r}\in V^{int}, \tag{2.23}$$

$$\mathbf{B}^{sca}(\mathbf{r}) = -\frac{ik}{\omega}\sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[p_{mn}\mathbf{M}_{mn}(k\mathbf{r}) + q_{mn}\mathbf{N}_{mn}(k\mathbf{r})\right] \quad \mathbf{r}\in V^{ext}, \tag{2.24}$$

where Eqs. (B.1)-(B.4) have been used.

The coefficients, a and b, for the incident wave Eq. (2.19) are known and given by Eqs. (B.15)-(B.17). Then, the coefficients c, d, p, and q for the internal and scattered waves can be found using the boundary conditions of Eqs. (1.40)-(1.43) along with the orthogonality relations found in App. B.

An important issue concerning the validity of the series expansions of Eqs. (2.20) and (2.21) is their convergence. It is the lack of convergence that essentially prevents application of the series-based method (eg. T-Matrix) to non-convex and highly irregularly shaped

particles. Care must be taken to ensure that enough terms are included in the expansion and that the solution is restricted from regions where its convergence is doubtful. More discussions on these issues are given in [32] and [2, p. 116-117].

Although the series-based solution for a particle's scattered wave does not play a major role in the development of the microphysical model, it does have value in that it provides an analytical framework on which one can make general investigations of fundamental aspects of the scattering process. For example, Ch. 3 will use the VSWF expansions to explicitly demonstrate that the total scattering, absorption, and extinction cross sections for an arbitrary particle are independent of distance. This will be an essential element in the explanation of the extinction paradox in Ch. 11 using the microphysical model.

2.3 Two Important Solutions

It is now possible to use the VIEs (2.9) and (2.17) to find explicit solutions to the wave equation. This will be done for two specific source distributions and will lead to expressions for the fields of spherical and planar waves. Both of these waves will be important for the microphysical model described in Ch. 4.

2.3.1 Spherical Wave

Consider the field resulting from Eq. (2.9) when the source function $\mathbf{j}$ corresponds to a point current element $\mathbf{J}^{pt}$. To express this explicitly, recall Eqs. (1.22) and (1.23),

$$\boldsymbol{\nabla} \times \mathbf{E}^{pt}(\mathbf{r}) - i\omega \mathbf{B}^{pt}(\mathbf{r}) = 0,$$

$$\boldsymbol{\nabla} \times \mathbf{B}^{pt}(\mathbf{r}) + i\omega \varepsilon_o \mu_o \mathbf{E}^{pt}(\mathbf{r}) = \mu_o \mathbf{J}^{pt}(\mathbf{r}).$$

Taking the curl of this first expression and substituting the second into the result gives

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}^{pt}(\mathbf{r}) - k^2 \mathbf{E}^{pt}(\mathbf{r}) = i\omega \mu_o \mathbf{J}^{pt}(\mathbf{r}). \tag{2.25}$$

Equating the right-hand side of this expression to that of Eq. (1.38) and recalling Eq. (2.9) shows that the Green's function method can be used to solve this equation for the field

produced by $\mathbf{J}^{pt}$ as

$$\mathbf{E}^{pt}(\mathbf{r}) = i\omega\mu_o \int_{\mathbb{R}^3} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}^{pt}(\mathbf{r}')\, \mathrm{d}\mathbf{r}'. \tag{2.26}$$

The point current density $\mathbf{J}^{pt}$ can be expressed in terms of its electric dipole moment $\mathbf{p}$ and the Dirac delta function as

$$\mathbf{J}^{pt}(\mathbf{r}) = -i\omega\delta(\mathbf{r} - \mathbf{r}_o)\mathbf{p}, \tag{2.27}$$

where $\mathbf{r}_o$ is the location of the point current [6, p. 371]. Combining this with Eq. (2.26) gives

$$\mathbf{E}^{pt}(\mathbf{r}) = \frac{k^2}{\varepsilon_o} \int_{V^{int}} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \cdot [\delta(\mathbf{r}' - \mathbf{r})\mathbf{p}]\, \mathrm{d}\mathbf{r}', \tag{2.28}$$

where $k = \omega\sqrt{\mu_o\varepsilon_o}$ has been used. Equation (2.28) can be readily evaluated to give

$$\mathbf{E}^{pt}(\mathbf{r}) = \frac{k^2}{\varepsilon_o} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}_o) \cdot \mathbf{p}. \tag{2.29}$$

Now take the location $\mathbf{r}_o$ of the point current to be at the origin. Then, using Eqs. (2.3) and (2.4), Eq. (2.29) becomes

$$\mathbf{E}^{pt}(\mathbf{r}) = \frac{k^2}{\varepsilon_o} \left(\overleftrightarrow{\mathbf{I}} + \frac{1}{k^2} \boldsymbol{\nabla} \otimes \boldsymbol{\nabla} \right) \cdot [g(r)\mathbf{p}], \tag{2.30}$$

where

$$g(r) = \frac{\exp(ikr)}{4\pi r}.$$

The partial derivatives can be evaluated in spherical polar coordinates following [33, p.79-80]. With the z-axis of the coordinate system oriented along $\mathbf{p}$, i.e. $\mathbf{p} = p\hat{\mathbf{z}}$,

$$\mathbf{E}^{pt}(\mathbf{r}) = \frac{k^2}{\varepsilon_o} p \left(\overleftrightarrow{\mathbf{I}} + \frac{1}{k^2} \boldsymbol{\nabla} \otimes \boldsymbol{\nabla} \right) \cdot [g(r)\hat{\mathbf{z}}], \tag{2.31}$$

and the dyadic product in Eq. (2.31) can be evaluated as

$$\begin{aligned} \boldsymbol{\nabla} \otimes \boldsymbol{\nabla} g(r) &= \frac{1}{r}\partial_r g(r) \overleftrightarrow{\mathbf{I}} + \left[\partial_r^2 g(r) - \frac{1}{r}\partial_r g(r) \right] \hat{\mathbf{r}} \otimes \hat{\mathbf{r}} \\ &= -\frac{\exp(ikr)}{4\pi r^3} \left[k^2 r^2 \hat{\mathbf{r}} \otimes \hat{\mathbf{r}} + (1 - ikr) \left(\overleftrightarrow{\mathbf{I}} - 3\hat{\mathbf{r}} \otimes \hat{\mathbf{r}} \right) \right]. \end{aligned} \tag{2.32}$$

This result combined with Eqs. (2.29) and (2.30) gives

$$\overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, 0) = \frac{\exp(ikr)}{4\pi r} \left[\left(\frac{1}{k^2 r^2} - \frac{i}{kr} \right) \left(3\hat{\mathbf{r}} \otimes \hat{\mathbf{r}} - \overleftrightarrow{\mathbf{I}} \right) + \overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}} \right]. \tag{2.33}$$

Substituting Eq. (2.33) into Eq. (2.29) and using relation

$$(\hat{\mathbf{r}} \otimes \hat{\mathbf{r}}) \cdot \mathbf{p} = \hat{\mathbf{r}} \cdot (\hat{\mathbf{r}} \cdot \mathbf{p}) \tag{2.34}$$

where $\hat{\mathbf{r}} = \mathbf{r}/r$ shows that

$$\begin{aligned} \mathbf{E}^{pt}(\mathbf{r}) &= \frac{k^2}{\varepsilon_o} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, 0) \cdot \mathbf{p} \\ &= \frac{k^2}{4\pi\varepsilon_o} \frac{\exp(ikr)}{r} \left\{ \left(\frac{1}{k^2 r^2} - \frac{i}{kr} \right) \left[3\left(\hat{\mathbf{r}} \cdot \mathbf{p}\right) \hat{\mathbf{r}} - \mathbf{p} \right] - \left(\hat{\mathbf{r}} \cdot \mathbf{p}\right) \hat{\mathbf{r}} + \mathbf{p} \right\}. \end{aligned} \tag{2.35}$$

Equation (2.35) is exactly the electric field produced by a time-harmonic electric dipole $\mathbf{p}$, c.f. [15, p. 411].

Now for the magnetic dyadic Green's function. Using Eq. (A.3) with Eq. (2.14) gives

$$\overleftrightarrow{\mathbf{G}}_m(\mathbf{r}, 0) = \frac{\exp(ikr)}{4\pi r} \left(ik - \frac{1}{r} \right) \left(\hat{\mathbf{r}} \times \overleftrightarrow{\mathbf{I}} \right). \tag{2.36}$$

Then, the magnetic field produced by the point current is

$$\begin{aligned} \mathbf{B}^{pt}(\mathbf{r}) &= -i\omega\mu_o \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}, 0) \cdot \mathbf{p} \\ &= \frac{k\omega\mu_o}{4\pi} \frac{\exp(ikr)}{r} \left(1 - \frac{1}{ikr} \right) \hat{\mathbf{r}} \times \mathbf{p}. \end{aligned} \tag{2.37}$$

The wave, given by the field pair $\mathbf{E}^{pt}$ and $\mathbf{B}^{pt}$ of Eqs. (2.35) and (2.37), has an important behavior at distances very far from the source compared to the wavelength $\lambda = 2\pi/k$,

$$\mathbf{E}^{pt}(\mathbf{r}) = \frac{k^2}{4\pi\varepsilon_o} \frac{\exp(ikr)}{r} \left[\mathbf{p} - \left(\hat{\mathbf{r}} \cdot \mathbf{p}\right) \hat{\mathbf{r}} \right] \quad kr \to \infty, \tag{2.38}$$

$$\mathbf{B}^{pt}(\mathbf{r}) = \frac{k\omega\mu_o}{4\pi} \frac{\exp(ikr)}{r} \hat{\mathbf{r}} \times \mathbf{p} \quad kr \to \infty, \tag{2.39}$$

where all terms in Eqs. (2.35) and (2.37) decaying faster than r^{-1} have been neglected. The wave described by these fields in the $kr \to \infty$ limit is a spherical wave, and is constant in amplitude and phase on spherical surfaces centered on the source location. This can be seen from the r dependence of the exponential terms in Eqs. (2.38) and (2.39). The wave is transverse and travels radially outward away from the source at the speed of light, c. One can verify that the fields are interrelated in the $kr \to \infty$ limit as

$$\mathbf{B}^{pt}(\mathbf{r}) = \frac{k}{\omega} \hat{\mathbf{r}} \times \mathbf{E}^{pt}(\mathbf{r}), \tag{2.40}$$

$$\mathbf{E}^{pt}(\mathbf{r}) = -\frac{\omega}{k}\hat{\mathbf{r}} \times \mathbf{B}^{pt}(\mathbf{r}). \tag{2.41}$$

2.3.2 Plane Wave

Now suppose that instead of a point current density there is a unidirectional surface-current density that covers the entire x-y plane passing through the origin, see Fig. (2.1). At each point on this plane, the surface current $\mathbf{K}$ is related to $\mathbf{J}^{pl}$ as

$$\mathbf{J}^{pl}(\mathbf{r}) = \mathbf{K}(\mathbf{r})\delta(z). \tag{2.42}$$

Substituting this into Eq. (2.26) gives

$$\mathbf{E}^{pl}(\mathbf{r}) = i\omega\mu_o \int_{\mathbb{R}^3} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \cdot \mathbf{K}(\mathbf{r}')\delta(z')\, \mathrm{d}\mathbf{r}',$$

which becomes

$$\mathbf{E}^{pl}(\mathbf{r}) = i\omega\mu_o \int_{\Pi_z} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \cdot \mathbf{K}(\mathbf{r}')\, \mathrm{d}\mathcal{S}', \tag{2.43}$$

where Π_z denotes the x-y plane through the origin, i.e. the plane with the z-axis parallel to its surface normal, see Fig. (6.2) in Ch. 6. A more general derivation of Eq. (2.43) will be given in Sec. (9.2). One can take the orientation of the coordinate system to be such that the direction of the current is parallel to the x-axis, then $\mathbf{K} = K\hat{\mathbf{x}}$. Moreover, one can assume that the observation point $\mathbf{r}$ resides on the positive z-axis at a distance z that is much greater[3] than λ, i.e. $\mathbf{r} = z\hat{\mathbf{z}}$ where $z/\lambda \gg 1$.

Now recall Eq. (2.33) and undo the restriction that that $\mathbf{r}'$ resides at the origin in the derivation of the Green's function. This can be done with the substitution $\mathbf{r} \to \mathfrak{r}$ in Eq. (2.33) where

$$\mathfrak{r} = \mathbf{r} - \mathbf{r}', \tag{2.44}$$

giving

$$\overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') = \frac{\exp(ik\mathfrak{r})}{4\pi\mathfrak{r}} \left[\left(\frac{1}{k^2\mathfrak{r}^2} - \frac{i}{k\mathfrak{r}} \right) \left(3\hat{\mathfrak{r}} \otimes \hat{\mathfrak{r}} - \overleftrightarrow{\mathbf{I}} \right) + \overleftrightarrow{\mathbf{I}} - \hat{\mathfrak{r}} \otimes \hat{\mathfrak{r}} \right]. \tag{2.45}$$

[3]Note that [25, p. 30-12] mentions that the end result of this calculation is the same regardless of the distance of $\mathbf{r}$ from the sheet.

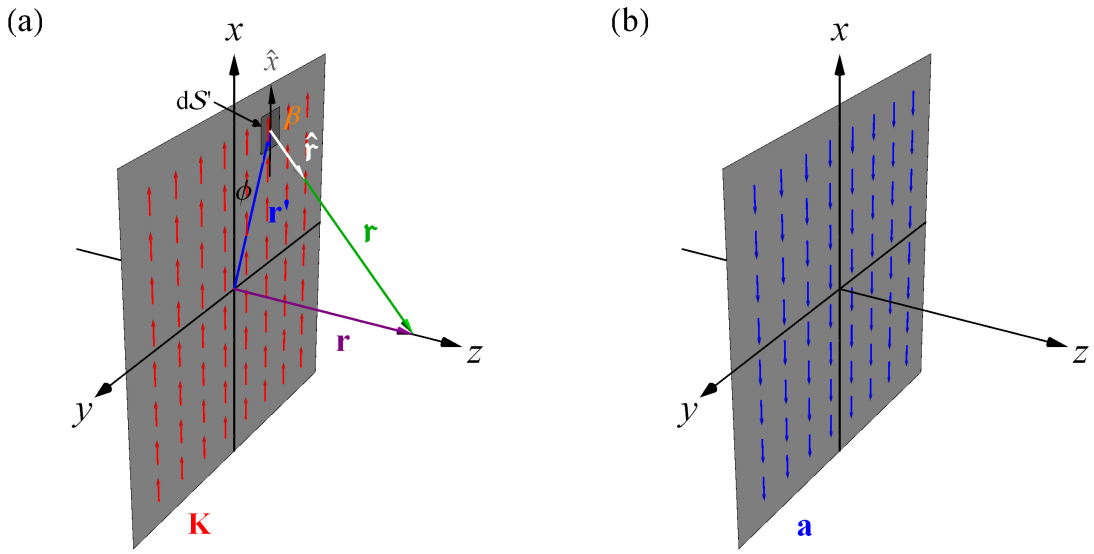

Figure 2.1: Planar surface current distribution $\mathbf{K}$, sketch (a), and corresponding surface dipole distribution $\mathbf{a}$, sketch (b).

Because of the assumed large distance between $\mathbf{r}$ and the current sheet, all terms in Eq. (2.45) decaying faster than $\mathfrak{r}^{-1}$ can be neglected, then

$$\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \frac{\exp(ik\mathfrak{r})}{4\pi\mathfrak{r}}\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathfrak{r}}\otimes\hat{\mathfrak{r}}\right). \tag{2.46}$$

Combining Eqs. (2.43) and (2.46) gives

$$\mathbf{E}^{pl}(\mathbf{r}) = i\omega\mu_o K\int_{\Pi_z}\frac{\exp(ik\mathfrak{r})}{4\pi\mathfrak{r}}\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathfrak{r}}\otimes\hat{\mathfrak{r}}\right)\cdot\hat{\mathbf{x}}\,\mathrm{d}\mathcal{S}'. \tag{2.47}$$

Equation (2.47) is simplified by introducing the angle β between $\mathbf{r}$ and $\mathbf{r}'$. This angle is shown in sketch (a) of Fig. (2.1). Since $\mathbf{r}'$ is in the x-y plane, one finds that

$$\begin{aligned}(\hat{\mathfrak{r}}\otimes\hat{\mathfrak{r}})\cdot\hat{\mathbf{x}} &= \hat{\mathfrak{r}}\,(\hat{\mathfrak{r}}\cdot\hat{\mathbf{x}})\\ &= \hat{\mathfrak{r}}\cos\left[\beta(\hat{\mathfrak{r}})\right].\end{aligned}$$

Using β in Eq. (2.47) and casting the resulting integral in Cartesian coordinates gives

$$\mathbf{E}^{pl}(z\hat{\mathbf{r}}) = \frac{i\omega\mu_o K}{4\pi}\iint_{-\infty}^{\infty}\frac{\exp\left[ik\left(x'^2+y'^2+z^2\right)^{1/2}\right]}{\left(x'^2+y'^2+z^2\right)^{1/2}}\left[\hat{\mathbf{x}} - \hat{\mathfrak{r}}\cos\beta(x',y')\right]\mathrm{d}x'\mathrm{d}y'. \tag{2.48}$$

Next, the integration in Eq. (2.48) is divided into two regions; the half of the plane with $x' \geqslant 0$ and the half with $x' < 0$. Inspection of sketch (a) in Fig. (2.1) reveals that β is bounded differently depending on which half of the plane the integration point is located in;

$$\beta(x', y') = \left[\frac{\pi}{2}, \pi\right], \quad x' \geqslant 0,$$
$$\beta(x', y') = \left[\frac{\pi}{2}, 0\right], \quad x' < 0.$$

This shows that the integral

$$\int_{-\infty}^{\infty}\int \frac{\exp\left[ik\left(x'^2 + y'^2 + z^2\right)^{1/2}\right]}{\left(x'^2 + y'^2 + z^2\right)^{1/2}} \cos\beta(x', y')\hat{\mathbf{t}}\,\mathrm{d}x'\mathrm{d}y' = 0,$$

since the term with the exponential in the integral is an even function whereas the remaining term $\cos\beta(x', y')\hat{\mathbf{t}}$ is an odd function as the *integration variables* cross from the $x' \geqslant 0$ to the $x' < 0$ half of the plane. Then Eq. (2.48) is

$$\mathbf{E}^{pl}(z\hat{\mathbf{r}}) = \frac{i\omega\mu_o K}{4\pi}\int_{-\infty}^{\infty}\int \frac{\exp\left[ik\left(x'^2 + y'^2 + z^2\right)^{1/2}\right]}{\left(x'^2 + y'^2 + z^2\right)^{1/2}}\,\mathrm{d}x'\mathrm{d}y'\,\hat{\mathbf{x}}. \tag{2.49}$$

This integral can be evaluated with cylindrical-polar coordinates,

$$\rho^2 = x'^2 + y'^2, \qquad \mathrm{d}x'\,\mathrm{d}y' = \rho\,\mathrm{d}\rho\,\mathrm{d}\phi,$$

giving

$$\mathbf{E}^{pl}(z\hat{\mathbf{r}}) = \frac{i\omega\mu_o K}{4\pi}\int_0^{\infty}\int_0^{2\pi} \frac{\exp\left[ik\left(\rho^2 + z^2\right)^{1/2}\right]}{\left(\rho^2 + z^2\right)^{1/2}}\rho\,\mathrm{d}\phi\,\mathrm{d}\rho\,\hat{\mathbf{x}}. \tag{2.50}$$

Evaluating this integral over the azimuthal angle and making the substitutions

$$\eta^2 = \rho^2 + z^2, \qquad \frac{\mathrm{d}\eta}{\mathrm{d}\rho'} = \frac{\rho}{(\rho^2 + z^2)^{1/2}},$$

gives

$$\mathbf{E}^{pl}(z\hat{\mathbf{r}}) = \frac{i\omega\mu_o K}{2}\int_z^{\infty}\exp(ik\eta)\,\mathrm{d}\eta\,\hat{\mathbf{x}}. \tag{2.51}$$

Further evaluation of Eq. (2.51) may seem problematic since it appears that

$$\int_z^{\infty}\exp(ik\eta)\,\mathrm{d}\eta = \frac{1}{ik}\left[\exp(i\infty) - \exp(ikz)\right],$$

is indefinite due to the term $\exp(i\infty)$. One can argue that this term is zero[4] by allowing k to have a small and positive imaginary part, k_{im} [35]. Let $k = k_{re} + ik_{im}$, then

$$\lim_{z\to\infty} \exp\left[i(k_{re} + ik_{im})z\right] = \lim_{z\to\infty} \exp(-k_{im}z)\exp(ik_{re}z) = 0,$$

which will be valid regardless of the value of k_{im} so long as z greatly exceeds $1/k_{im}$. With this result, one now finally has the expression for the field produced by the surface current,

$$\mathbf{E}^{pl}(z\hat{\mathbf{r}}) = -\frac{k}{2\omega\varepsilon_o} K \exp(ikz)\,\hat{\mathbf{x}}. \tag{2.52}$$

Just as the wave produced by a point current source is expressed in Eqs. (2.35) and (2.37) in terms of its corresponding electric dipole moment $\mathbf{p}$, Eq. (2.52) can be expressed in terms of an electric surface dipole moment $\mathbf{a}$. To do this, equate Eq. (2.27) and (2.42), giving

$$\mathbf{K}(\mathbf{r}')\delta(z') = -i\omega\delta(x - x')\delta(y - y')\delta(z')\mathbf{p},$$

then,

$$\mathbf{K}(\mathbf{r}) = -i\omega\delta(x - x')\delta(y - y')\mathbf{a}. \tag{2.53}$$

Notice the resemblance to Eq. (2.27) and that the dipole moment is directed opposite to the current and shifted in phase by $\pi/2$, see sketch (b) in Fig. (2.1). From Eq. (2.53), the current magnitude is $K = -i\omega a$, then

$$\mathbf{E}^{pl}(\mathbf{r}) = \frac{ika}{2\varepsilon_o} \exp(ikz)\,\hat{\mathbf{a}}. \tag{2.54}$$

Using Eq. (1.29), the corresponding magnetic field is

$$\mathbf{B}^{pl}(\mathbf{r}) = \frac{i\omega a\mu_o}{2} \exp(ikz)\,(\hat{\mathbf{z}} \times \hat{\mathbf{a}})\,. \tag{2.55}$$

By comparison to [1, Sec. 2.5], one can see that Eqs. (2.54) and (2.55) are specific examples of a more general solution for the fields due to this source distribution,

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}_o \exp(ikr\hat{\mathbf{n}} \cdot \hat{\mathbf{r}}), \tag{2.56}$$

[4] An alternative way to show that this is zero is based on "physical arguments" in [34, p. 30-11].

$$\mathbf{B}(\mathbf{r}) = \frac{k}{\omega}\left(\hat{\mathbf{n}} \times \mathbf{E}_o\right) \exp(ikr\hat{\mathbf{n}} \cdot \hat{\mathbf{r}}). \tag{2.57}$$

The vector $\hat{\mathbf{n}}$ in Eqs. (2.56) and (2.57) describes the propagation direction of the wave, whereas $\mathbf{E}_o$ describes the amplitude of the wave and is directed along $\mathbf{a}$. One can now see that the fields of Eqs. (2.56) and (2.57) describe a homogeneous plane wave[5] with parallel planar-surfaces of constant phase and amplitude. Comparison of Eqs. (2.56) and (2.57) shows that the fields are related by

$$\mathbf{B}(\mathbf{r}) = \frac{k}{\omega}\hat{\mathbf{n}} \times \mathbf{E}(\mathbf{r}), \tag{2.58}$$

$$\mathbf{E}(\mathbf{r}) = -\frac{\omega}{k}\hat{\mathbf{n}} \times \mathbf{B}(\mathbf{r}). \tag{2.59}$$

[5]It is curious to note that Eqs. (2.56) and (2.57) can be derived by assuming that there are no sources anywhere in space [1, Sec. 2.5]. Physically, this seems to be inconsistent in the context of the above calculations, which find the same plane wave solution but do so for an infinite source distribution. Rather than dwell on the significance of this issue, this work will take the position that the a plane wave is created by some time-harmonic source that is very far from the particle and is independent of its presence.

Chapter 3

General Derivation of the Total Electromagnetic Cross Sections

> "[*the conservation of energy*] *is not a description of a mechanism, or anything concrete; it is just a strange fact that we can calculate some number and when we finish watching nature go through her tricks and calculate the number again, it is the same.*"
>
> -R. P. Feynman [34, p. 4-1]

The total scattering, absorption, and extinction cross sections, denoted C^{sca}, C^{abs}, and C^{ext}, respectively, are integrated quantities that collectively account for the redistribution and conservation of energy in the scattering process. Consequently, the cross sections are important quantities with regard to understanding how the physical properties of a particle relate to its ability to achieve the redistribution of the energy content of an incident wave. Moreover, the cross sections can be measured, and this gives them more than theoretical significance. A great deal of attention will be given to the cross sections later on, and because of this, it is important to establish general and exact expressions for them here.

Typically, one calculates the cross sections in the particle's far-field zone. This is done partly because the mathematical form of the fields are substantially simpler there, and because most cross section measurements are performed far enough from the particle that the far-field approximation can be quite good. Now recall that the particle's external medium is vacuum, and hence is nonabsorbent. Since the cross sections account for the conservation

of energy, one would then expect that expressions for them should be independent of the distance from the particle at which they are calculated. The following will demonstrate how exact expressions for the total cross sections can be found that are independent of distance from the particle. This is done by explicitly evaluating the integrals giving the cross sections in a particle's near-field zone. The exact equivalence of the resulting expressions to those that are derived in the far-field zone proves the distance independence. Numerical calculations for spherical particles are presented that verify the equivalence of the cross sections when calculated in the near and far-field zones and help illustrate the physical significance of this result.

One should note that this work is not the first to calculate a particle's cross sections in the near-field. For example, work by [36, p. 70-71] and [37, Sec. 4] mentions that C^{sca} and C^{abs} can be found using a surface, like $\mathcal{S}_{en}$ described below, that is of arbitrary size, but the details of the calculation are not given. Also, [38, 39] find near-field expressions for the cross sections for a particle embedded in an *absorbing* medium. Further relevant work related to particles in an absorbing medium can be found in [40, 41] and references therein. To the author's knowledge, the following calculations are the first to consider an *arbitrary* particle and formulate and evaluate exact expressions for the total cross sections valid at *any* distance from it.

3.1 Derivation of the Cross Sections

Consider the arrangement shown in Fig. (3.1). Surrounding the particle is its smallest circumscribing sphere $\mathcal{S}_{sc}$ of radius R_{sc}. The purpose of this sphere is to define the minimum distance from the origin beyond which one is guaranteed that the VSWF expansions of Eqs. (2.19), (2.21), (2.22), and (2.24) will converge. Also enclosing the particle, and $\mathcal{S}_{sc}$, is another spherical surface $\mathcal{S}_{en}$ of radius R_{en}. This sphere will be used as the closed surface over which the fields will be integrated in Eqs. (3.5) and (3.6) below to yield expressions for the total cross sections.

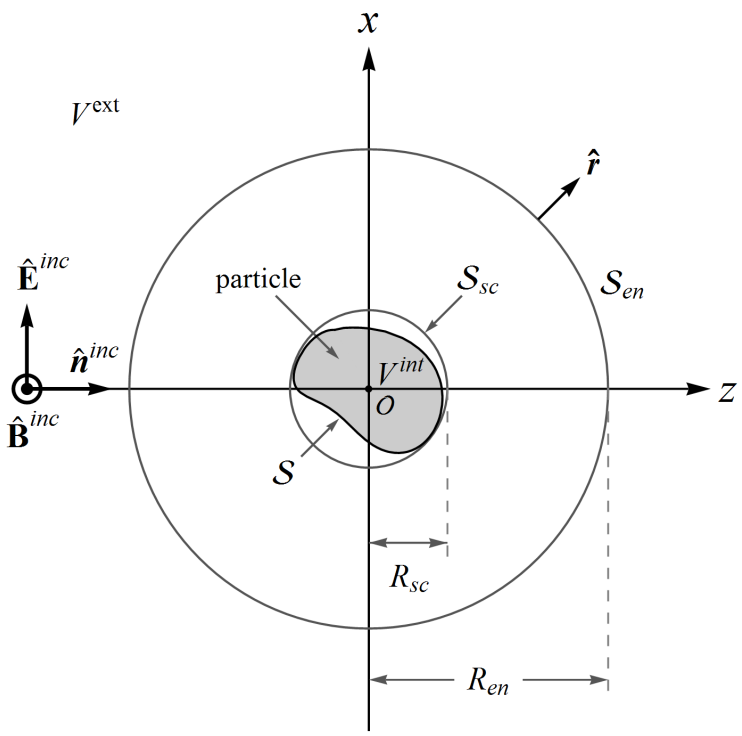

Figure 3.1: Arrangement used to derive the cross sections for an arbitrary particle.

The wave incident on the particle is a linearly-polarized plane wave traveling along the $\hat{\mathbf{n}}^{inc}$ direction with a wavenumber $k = 2\pi/\lambda$, where λ is the vacuum wavelength. This is the same plane wave presented in Eqs. (2.54) and (2.55) with $\hat{\mathbf{n}}$ in those expressions replaced by $\hat{\mathbf{n}}^{inc}$ here to avoid notational confusion with surface normals. The fields of the incident wave are

$$\mathbf{E}^{inc}(\mathbf{r}) = \mathbf{E}_o^{inc} \exp(ikr\hat{\mathbf{n}}^{inc} \cdot \hat{\mathbf{r}}), \tag{3.1}$$

and

$$\mathbf{B}^{inc}(\mathbf{r}) = \frac{k}{\omega}\, \hat{\mathbf{n}}^{inc} \times \mathbf{E}^{inc}(\mathbf{r})\,, \tag{3.2}$$

respectively, where $\mathbf{E}_o^{inc}$ is a constant vector describing the amplitude and polarization direction of the wave. Throughout the remainder of this work, a particle's incident wave will be assumed to be the plane wave described by these fields. The reason that a plane wave is chosen is partly because of its simple mathematical form and because it is often a good approximation for the light emitted by a typical laser used in simple[1] scattering

[1]This would exclude ultra-short laser pulse illumination, where "short" is quantified by the degree to

measurements.

As discussed in Sec. 1.7, the presence of the particle in the incident wave establishes a new (total) wave that can be mathematically regarded as the superposition of the incident and scattered waves. The total electric and magnetic fields outside of the particle are then

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \mathbf{E}^{sca}(\mathbf{r}), \tag{3.3}$$

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}^{inc}(\mathbf{r}) + \mathbf{B}^{sca}(\mathbf{r}). \tag{3.4}$$

These fields can be expanded into series of vector spherical wave functions (VSWFs) as given by Eqs. (2.19), (2.21), (2.22), and (2.24),

$$\mathbf{E}^{inc}(\mathbf{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[a_{mn}\mathrm{Rg}\mathbf{M}_{mn}(k\mathbf{r}) + b_{mn}\mathrm{Rg}\mathbf{N}_{mn}(k\mathbf{r})\right] \quad \mathbf{r}\in V^{int}\cup V^{ext},$$

$$\mathbf{E}^{sca}(\mathbf{r}) = \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\left[p_{mn}\mathbf{M}_{mn}(k\mathbf{r}) + q_{mn}\mathbf{N}_{mn}(k\mathbf{r})\right] \quad \mathbf{r}\in V^{ext}.$$

The properties of the VSWFs $\mathbf{M}$, $\mathbf{N}$, $\mathrm{Rg}\mathbf{M}$, and $\mathrm{Rg}\mathbf{N}$ in Eqs. (2.19) and (2.21) needed in the following calculations are given in App. B. A more complete collection of VSWF properties and their definitions is given in [2, App. C]. Also recall the discussions regarding the properties of the VSWFs given in Sec. 2.2.

Expressions for the total cross sections are obtained by integrating the component of the total flow of electromagnetic energy passing through a closed surface enclosing the particle. The spherical surface $\mathcal{S}_{en}$ will serve as this closed surface, where the flow of electromagnetic energy is given by the time-averaged Poynting vector $\langle\mathbf{S}\rangle_t$ of Eq. (1.45),

$$\langle\mathbf{S}(\mathbf{r})\rangle_t = \frac{1}{2\mu_o}\,\mathrm{Re}\Big\{\mathbf{E}(\mathbf{r})\times\big[\mathbf{B}(\mathbf{r})\big]^*\Big\}.$$

Using Eqs. (3.3), (3.4), and (1.45) and noting that the radial unit vector $\hat{\mathbf{r}}$ is also the outward surface normal for $\mathcal{S}_{en}$, one finds the following expressions for the total extinction

which the pulse's bandwidth can be considered monochromatic. Also excluded would be tightly focused beams.

and scattering cross sections in terms of the fields explicitly,

$$C^{ext} = -\frac{1}{2\mu_o I^{inc}} \mathrm{Re} \oint\limits_{\mathcal{S}_{en}} \left\{ \mathbf{E}^{inc}(\mathbf{r}) \times \left[\mathbf{B}^{sca}(\mathbf{r})\right]^* + \mathbf{E}^{sca}(\mathbf{r}) \times \left[\mathbf{B}^{inc}(\mathbf{r})\right]^* \right\} \cdot \hat{\mathbf{r}} \, \mathrm{d}S, \tag{3.5}$$

$$C^{sca} = \frac{1}{2\mu_o I^{inc}} \mathrm{Re} \oint\limits_{\mathcal{S}_{en}} \left\{ \mathbf{E}^{sca}(\mathbf{r}) \times \left[\mathbf{B}^{sca}(\mathbf{r})\right]^* \right\} \cdot \hat{\mathbf{r}} \, \mathrm{d}S, \tag{3.6}$$

where $I^{inc} = (1/2)\sqrt{\epsilon_o/\mu_o}|\mathbf{E}_o^{inc}|^2$ is the intensity of the incident wave. Chapter 10 will give more detail concerning the derivation and physical meaning of Eqs. (3.5) and (3.6).

The size of $\mathcal{S}_{en}$ in Eqs. (3.5) and (3.6), given by R_{en}, is arbitrary provided that $R_{en} \geqslant R_{sc}$. Consequently, one usually regards R_{en} as being large enough that $\mathcal{S}_{en}$ resides in the particle's far-field zone. This permits approximation of the scattered fields of Eqs. (3.3) and (3.4) as outward-traveling transverse spherical waves and justifies use of the optical theorem; doing so simplifies the evaluation of Eqs. (3.5) and (3.6), see [42, 43].

In the following sections, the VSWF expansions of Eqs. (2.19), (2.21), (2.22), and (2.24) are combined with Eqs. (3.5) and (3.6) to arrive at general expressions for C^{ext} and C^{sca}. This is done specifically *without* any assumption regarding the size of $\mathcal{S}_{en}$ and *without* invoking the far-field approximation. One will see that the calculation of C^{ext} and C^{sca}, while somewhat more mathematically involved, yields expressions for the cross sections that are *exactly* identical to those that would result from use of the far-field approximation.

3.1.1 Extinction Cross Section

Combining Eqs. (2.19), (2.21), (2.22), and (2.24) with Eq. (3.5), grouping terms with common expansion coefficients, and using the vector identity $(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a}=\mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$ gives,

$$
\begin{aligned}
C^{ext} = -c_o \mathrm{Re} \oint_{\mathcal{S}_{en}} i \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \sum_{n'=1}^{\infty} \sum_{m'=-n'}^{n'} \Big\{ & a_{m'n'}(p_{mn})^* \mathbf{N}^*_{mn}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathrm{Rg}\mathbf{M}_{m'n'}(k\mathbf{r})\right] \\
& + (a_{m'n'})^* p_{mn} \mathrm{Rg}\mathbf{N}^*_{m'n'}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathbf{M}_{mn}(k\mathbf{r})\right] \\
& + b_{m'n'}(q_{mn})^* \mathbf{M}^*_{mn}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathrm{Rg}\mathbf{N}_{m'n'}(k\mathbf{r})\right] \\
& + (b_{m'n'})^* p_{mn} \mathrm{Rg}\mathbf{M}^*_{m'n'}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathbf{N}_{mn}(k\mathbf{r})\right] \\
& + b_{m'n'}(p_{mn})^* \mathbf{N}^*_{mn}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathrm{Rg}\mathbf{N}_{m'n'}(k\mathbf{r})\right] \\
& + (a_{m'n'})^* q_{mn} \mathrm{Rg}\mathbf{N}^*_{m'n'}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathbf{N}_{mn}(k\mathbf{r})\right] \\
& + a_{m'n'}(q_{mn})^* \mathbf{M}^*_{mn}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathrm{Rg}\mathbf{M}_{m'n'}(k\mathbf{r})\right] \\
& + (b_{m'n'})^* p_{mn} \mathrm{Rg}\mathbf{M}^*_{m'n'}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathbf{M}_{mn}(k\mathbf{r})\right] \Big\} \, \mathrm{d}S, \quad (3.7)
\end{aligned}
$$

where $c_o = k/2\omega\mu_o I^{inc}$. Next, using appendix Eqs. (B.5)-(B.13) shows that the following terms in Eq. (3.7) above are zero;

$$\oint_{\mathcal{S}_{en}} \mathbf{M}^*_{mn}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathrm{Rg}\mathbf{M}_{m'n'}(k\mathbf{r})\right] \mathrm{d}S = 0, \quad (3.8)$$

$$\oint_{\mathcal{S}_{en}} \mathrm{Rg}\mathbf{M}^*_{m'n'}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathbf{M}_{mn}(k\mathbf{r})\right] \mathrm{d}S = 0, \quad (3.9)$$

$$\oint_{\mathcal{S}_{en}} \mathbf{N}^*_{mn}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathrm{Rg}\mathbf{N}_{m'n'}(k\mathbf{r})\right] \mathrm{d}S = 0, \quad (3.10)$$

$$\oint_{\mathcal{S}_{en}} \mathrm{Rg}\mathbf{N}^*_{m'n'}(k\mathbf{r}) \cdot \left[\hat{\mathbf{r}} \times \mathbf{N}_{mn}(k\mathbf{r})\right] \mathrm{d}S = 0. \quad (3.11)$$

Using the orthogonality relation of Eq. (B.14) along with Eqs. (B.5)-(B.9) on the remaining terms in Eq. (3.7) eventually gives

$$\begin{aligned} C^{ext} = -c_o R_{en}^2 \mathrm{Re} \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \Big\{ & i a_{mn} (p_{mn})^* \frac{1}{kR_{en}} \Big[kR_{en} \big[h_n^{(1)}(kR_{en}) \big]^* \Big]' j_n(kR_{en}) \\ & + i (a_{mn})^* p_{mn} \frac{1}{kR_{en}} \big[kR_{en} j_n(kR_{en}) \big]' h_n^{(1)}(kR_{en}) \\ & - i b_{mn} (q_{mn})^* \frac{1}{kR_{en}} \big[kR_{en} j_n(kR_{en}) \big]' \big[h_n^{(1)}(kR_{en}) \big]^* \\ & - i (b_{mn})^* q_{mn} \frac{1}{kR_{en}} \big[kR_{en} h_n^{(1)}(kR_{en}) \big]' j_n(kR_{en}) \Big\}, \end{aligned} \tag{3.12}$$

where $\big[kR_{en} h_n^{(1)}(kR_{en})\big]'$ denotes differentiation with respect to kR_{en}. This expression is simplified using the linear independence of the spherical Bessel and Hankel functions via the Wronskian relation [15, p. 427]

$$W_1 = \big[kR_{en} j_n(kR_{en})\big]' h_n^{(1)}(kR_{en}) - \big[kR_{en} h_n^{(1)}(kR_{en})\big]' j_n(kR_{en}) = -\frac{i}{kR_{en}}. \tag{3.13}$$

With Eq. (3.13), and keeping in mind that the spherical Bessel functions of the first kind j_n are real-valued since their argument is real-valued, Eq. (3.12) can be reduced to

$$C^{ext} = -\frac{1}{k^2 |\mathbf{E}_o^{inc}|^2} \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \mathrm{Re}\big[a_{mn} (p_{mn})^* + b_{mn} (q_{mn})^* \big]. \tag{3.14}$$

One can verify that Eq. (3.14) is exactly the same expression for C^{ext} obtained from use of the far-field approximation for the scattered wave in combination with the optical theorem, eg. see [2, Eq. (5.18a)] and [44, p. 60-61].

3.1.2 Scattering Cross Section

Combining Eqs. (2.21) and (2.24) with (3.6) and using the vector identity $(\mathbf{b}\times\mathbf{c})\cdot\mathbf{a} = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$ gives for the scattering cross section

$$\begin{aligned} C^{sca} = c_o \mathrm{Re} \oint_{\mathcal{S}_{en}} i \sum_{n=1}^{\infty}\sum_{m=-n}^{n}\sum_{n'=1}^{\infty}\sum_{m'=-n'}^{n'} \Big\{ & p_{mn}(p_{m'n'})^* \mathbf{N}^*_{m'n'}(k\mathbf{r})\cdot\big[\hat{\mathbf{r}}\times\mathbf{M}_{mn}(k\mathbf{r})\big] \\ & + p_{mn}(q_{m'n'})^* \mathbf{M}^*_{m'n'}(k\mathbf{r})\cdot\big[\hat{\mathbf{r}}\times\mathbf{M}_{mn}(k\mathbf{r})\big] \\ & + (p_{m'n'})^* q_{mn}\mathbf{N}^*_{m'n'}(k\mathbf{r})\cdot\big[\hat{\mathbf{r}}\times\mathbf{N}_{mn}(k\mathbf{r})\big] \\ & + q_{mn}(q_{m'n'})^*\mathbf{M}^*_{m'n'}(k\mathbf{r})\cdot\big[\hat{\mathbf{r}}\times\mathbf{N}_{mn}(k\mathbf{r})\big]\Big\}\,\mathrm{d}S. \end{aligned} \tag{3.15}$$

From appendix Eqs. (B.5), (B.7), and (B.9)-(B.13), one finds that

$$\oint_{\mathcal{S}_{en}} \mathbf{M}^*_{m'n'}(k\mathbf{r})\cdot\big[\hat{\mathbf{r}}\times\mathbf{M}_{mn}(k\mathbf{r})\big]\,\mathrm{d}S = 0, \tag{3.16}$$

$$\oint_{\mathcal{S}_{en}} \mathbf{N}^*_{m'n'}(k\mathbf{r})\cdot\big[\hat{\mathbf{r}}\times\mathbf{N}_{mn}(k\mathbf{r})\big]\,\mathrm{d}S = 0, \tag{3.17}$$

which reduces Eq. (3.15) to

$$\begin{aligned} C^{sca} = \frac{c_o R_{en}}{k}\sum_{n=1}^{\infty}\sum_{m=-n}^{n} \mathrm{Re}\Big\{ & i|p_{mn}|^2\Big[kR_{en}\big[h_n^{(1)}(kR_{en})\big]^*\Big]' h_n^{(1)}(kR_{en}) \\ & - i|q_{mn}|^2\big[kR_{en}h_n^{(1)}(kR_{en})\big]'\big[h_n^{(1)}(kR_{en})\big]^*\Big\}. \end{aligned} \tag{3.18}$$

To simplify Eq. (3.18), first recall that $h_n^{(1)}$ and $h_n^{(2)}$ are related to j_n and y_n as

$$h_n^{(1)}(kr) = j_n(kr) + iy_n(kr), \tag{3.19}$$

$$h_n^{(2)}(kr) = j_n(kr) - iy_n(kr). \tag{3.20}$$

The argument of the spherical Bessel functions j_n and y_n is real-valued outside of the particle since the external medium is nonabsorbent. Consequently, the functions are also real-valued. Then, Eqs. (3.19) and (3.20) show that

$$h_n^{(2)}(kR_{en}) = \big[h_n^{(1)}(kR_{en})\big]^*. \tag{3.21}$$

Using Eq. (3.21) with Eq. (3.18) gives

$$C^{sca} = \frac{c_o R_{en}}{k} \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \Bigg[|p_{mn}|^2 \, \mathrm{Re}\Big\{ i\big[kR_{en} h_n^{(2)}(kR_{en})\big]' h_n^{(1)}(kR_{en}) \Big\} - |q_{mn}|^2 \, \mathrm{Re}\Big\{ i\big[kR_{en} h_n^{(1)}(kR_{en})\big]' h_n^{(2)}(kR_{en}) \Big\} \Bigg]. \quad (3.22)$$

The Wronskian relation for the spherical Bessel functions of the first and second kind is [45, p. 437]

$$W_2 = j_n(kR_{en}) y_n'(kR_{en}) - y_n(kR_{en}) j_n'(kR_{en}) = \frac{1}{(kR_{en})^2}. \quad (3.23)$$

By using Eq. (3.23) and by neglecting all terms that are pure imaginary due to the presence of the Re{...} filter, one eventually finds that the first term in Eq. (3.22) is

$$\mathrm{Re}\Big\{ i\big[kR_{en} h_n^{(2)}(kR_{en})\big]' h_n^{(1)}(kR_{en}) \Big\} = \frac{1}{kR_{en}}, \quad (3.24)$$

and that the second term is

$$\mathrm{Re}\Big\{ -i\big[kR_{en} h_n^{(1)}(kR_{en})\big]' h_n^{(2)}(kR_{en}) \Big\} = -\frac{1}{kR_{en}}. \quad (3.25)$$

With Eqs. (3.24) and (3.25), the total scattering cross section of Eq. (3.22) becomes

$$C^{sca} = \frac{1}{k^2 |\mathbf{E}_o^{inc}|^2} \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \Big\{ |p_{mn}|^2 + |q_{mn}|^2 \Big\}, \quad (3.26)$$

which is exactly what is expected from use of the far-field approximation for the scattered wave, eg. see [2, Eq. (5.18b)].

3.1.3 Absorption Cross Section

From energy conservation considerations, one finds that the extinction, scattering, and absorption cross sections are related by

$$C^{ext} = C^{sca} + C^{abs},$$

e.g., see Ch. 10. With this relationship, Eqs. (3.14) and (3.26) can be combined to yield the total absorption cross section,

$$C^{abs} = -\frac{1}{k^2 |\mathbf{E}_o^{inc}|^2} \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \Big\{ \mathrm{Re}\big[a_{mn} (p_{mn})^* + b_{mn} (q_{mn})^* \big] + |p_{mn}|^2 + |q_{mn}|^2 \Big\}. \quad (3.27)$$

3.2 Numerical Examples

Equations (3.14), (3.26), and (3.27) are *exact* expressions. No approximations, specifically the far-field approximation, are needed for their derivation. The expressions are valid for a particle of any shape and composition provided that it resides in a nonabsorbent medium and is illuminated by a linearly-polarized electromagnetic plane wave. In what follows numerical examples involving spherical particles will be presented that demonstrate implications of the above results.

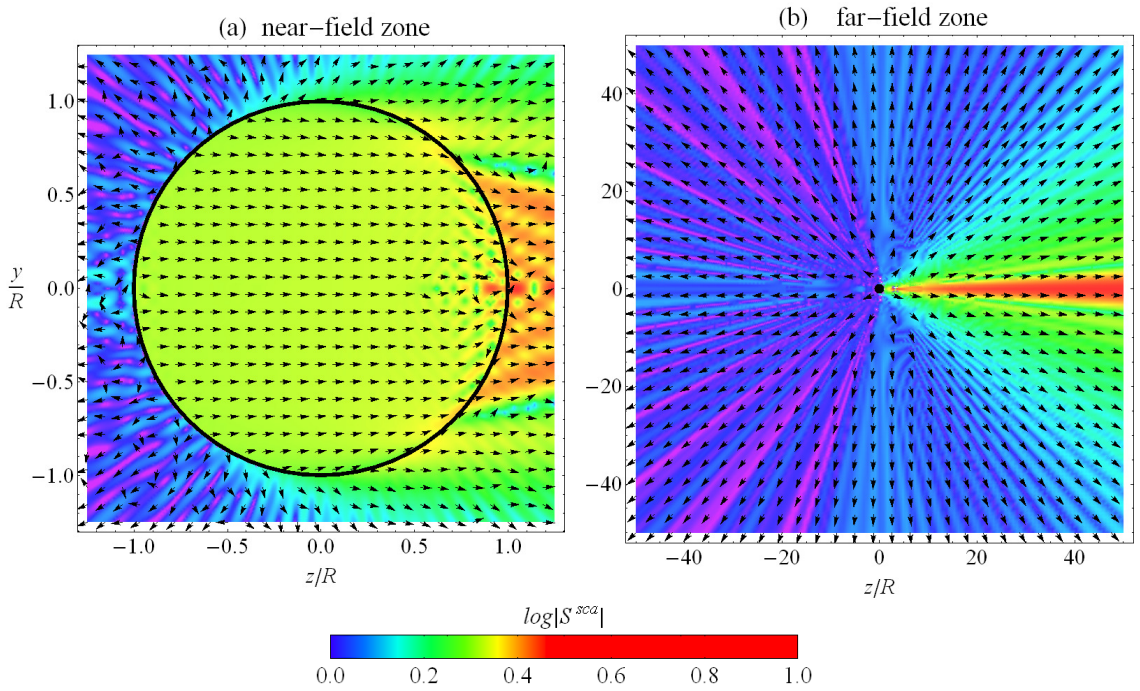

Figure 3.2: Time-averaged energy flow due to a spherical particle, where $kR = 50$ and $m = 1.10 + 0i$. The incident wave travels along the z-axis from left to right and is polarized along the x-axis, normal to the page. Plot (a) shows the energy flow $\langle \mathbf{S}^{sca} \rangle_t$ of Eq. (3.28) in the sphere's near field zone outside the particle, and the total energy flow of Eq. (1.45) inside the particle. Plot (b) shows $r^2 \langle \mathbf{S}^{sca} \rangle_t$ in the particle's far-field zone; the flow is multiplied by r^2 to help reveal its angular structure. Both plots show the y-z plane through the origin, which slices the sphere along its equator. The color shades indicate the magnitude of the energy flow in log scale with the largest magnitude normalized to unity, i.e. $\log(1.0) = 0.0$.

Consider the calculation of C^{sca} for a spherical particle with $kR = 50$ and $m = 1.10 + 0i$. The scattering cross section is determined by the term in the total energy flow of Eq. (1.45)

involving only the fields of the scattered wave,

$$\langle \mathbf{S}^{sca}(\mathbf{r})\rangle_t = \frac{1}{2\mu_o}\,\mathrm{Re}\Big\{\mathbf{E}^{sca}(\mathbf{r})\times\big[\mathbf{B}^{sca}(\mathbf{r})\big]^*\Big\}, \tag{3.28}$$

see [2, Sec. 2.8]. To calculate C^{sca} from Eq. (3.28), the radial component of the energy flow is integrated over $\mathcal{S}_{en}$. Recall that the radius of this surface R_{en} is arbitrary; $\mathcal{S}_{en}$ can be either in the particle's near or far-field zone.

Figure (3.2) shows the energy flow of Eq. (3.28) for the spherical particle in the near and far-field zones, plots (a) and (b), respectively. In plot (a), the energy flow that is shown on the inside of the particle is the total flow of Eq. (1.45) given by the internal fields. In plot (b), the energy flow is multiplied by r^2 to reveal its angular structure. The incident wave propagates along the z-axis from left to right and is polarized along the x-axis. All field quantities are calculated from the Mie solution to the Maxwell equations following [37, Ch. 4].

One can see that the plots in Fig. (3.2) show vastly different behavior for the energy flow between the two zones. This difference is no surprise. The scattered wave in the near-field zone can be complicated in form; it can have a longitudinal component and nonspherical surfaces of constant phase and amplitude. In the far-field zone however, the scattered wave becomes a transverse spherical wave with spherical surfaces of constant phase and amplitude [43]. This gives the far-field energy flow its simple radial-direction and angular magnitude-dependence. Despite the dramatic differences apparent in Fig. (3.2), the exact same cross sections are obtained from the energy flows.

To demonstrate numerically the distance independence of the cross sections, the extinction efficiency Q^{ext} for several spherical particles is calculated in both the near and far-field zones and presented in Table 3.1. To do this, the integral in Eq. (3.5) is numerically evaluated over $\mathcal{S}_{en}$ when $R_{en} = 1.001R$. Provided that $kR >> 1$, as it is in Table 3.1 below, this size of $\mathcal{S}_{en}$ places the surface in the particle's extreme near-field zone. The resulting value for the cross section will be denoted the near-field extinction cross section C^{ext}_{near}. From C^{ext}_{near}, the near-field extinction efficiency follows as $Q^{ext}_{near} = C^{ext}_{near}/\pi R^2$. For comparison, the extinc-

tion cross section in the particle's far-field zone, denoted C^{ext}_{far}, is calculated using the optical theorem, e.g. see [2, p. 57]. This gives a far-field extinction efficiency $Q^{ext}_{far} = C^{ext}_{far}/\pi R^2$.

kR	m	Q^{ext}_{near}	Q^{ext}_{far}
50.00	$1.100 + 0.000i$	2.449	2.446
50.00	$1.330 + 1.000 \times 10^{-6}i$	1.980	1.982
25.00	$1.750 + 0.440i$	2.224	2.221
25.00	$1.010 + 1.000i$	2.233	2.231
10.00	$10.00 + 0.000i$	2.106	2.104

Table 3.1: Values of Q^{ext} for spherical particles in the near and far-field.

Table 3.1 shows values for Q^{ext}_{near} and Q^{ext}_{far} for spherical particles of vastly different size kR and refractive index m. One can see that the efficiency factors are equivalent to 0.1%. The disagreement between Q^{ext}_{near} and Q^{ext}_{far} is due to the discretization of the surface integral in Eq. (3.5) that is required to evaluate Q^{ext}_{near} numerically.

3.3 Comments

As mentioned above, one typically calculates a particle's total cross section using the far-field zone approximation for the scattered wave. For example, in the case of the extinction cross section, a far-field scattering amplitude is found that is independent of distance from the particle. The optical theorem is then used to relate the extinction cross section to the value of the scattering amplitude evaluated in the exact forward direction. This will be described in detail in Ch. 10. The result is an exact expression matching Eq. (3.14), despite the use of an approximate form for the scattered wave. One might then wonder how an approximation for the scattered wave, i.e. the far-field approximation, can result in an exact expression for the cross section. The limit $kR_{en} \to \infty$ must be assumed in using the optical theorem, e.g. see [2, App. A] and Ch. 10. In this *limit*, the far-field zone approximation becomes exact and hence is capable of producing an exact expression for C^{ext}. The point here is that the ability to find exact expressions for the total cross sections using the far-field approximation

does not necessarily restrict the validity of the resulting expressions to only the far-field zone.

Another integral quantity that is often of interest is the asymmetry parameter g, which describes the overall tendency for a particle to scatter into the forward or backward directions, see [2, p. 60]. Like the total cross sections, the asymmetry parameter is typically calculated in a particle's far-field zone owing to the mathematical convenience of the far-field approximation. However, since g is ultimately derived from the same field quantities that yield C^{sca}, i.e., the scattered fields, it should be possible to carry out a calculation for g similar to that for C^{sca} above that yields g for any distance from the particle. Work by [46] shows how this can be done in the context of the same VSWF formalism used above.

Notice that if the particle were to reside in an absorbing medium, the total cross sections *would* depend on the distance at which they are calculated, eg. see [38]. In the context of the calculations in Sec. 3.1, an absorbing external medium would make the wave number k complex valued. Consequently, the spherical Bessel functions would also become complex valued and one would not be able to use the $\mathrm{Re}\{\ldots\}$ filter in combination with the Wronskians of Eqs. (3.13) and (3.23) to remove the R_{en} dependence from Eqs. (3.12) and (3.22).

The results in this chapter are not solely of mathematical interest. In hindsight, one could expect that there must be distance independence to the total cross sections since, collectively, the cross sections are a statement of energy conservation. As such, they cannot rely on some assumed (far field) distance from the particle. One can extend this mindset to other phenomena in electromagnetic theory. For example, Ch. 11 shows that the extinction cross section for a large particle approaches a value of $C^{ext} \to 2G$ as the particle's size becomes large compared to the vacuum wavelength. Here G is the particle's geometric projection into the forward direction, and this result is known as the extinction paradox. The popular explanations for this paradox are implicitly restricted to the far-field zone only, e.g. [47, Sec. 8.22] and [48]. The essence of the results above demonstrate that explanations

of phenomena like the extinction paradox must be able to apply at all distances from the particle, not just in the far-field zone.

Chapter 4

Microphysical Model

> *"The most dramatic movements in the development of physics are those in which great syntheses take place, where phenomena which previously had appeared to be different are suddenly discovered to be but different aspects of the same thing."*
>
> -R. P. Feynman [34, p. 28-1]

Here, the internal electric field inside of a differential volume element in the particle will be shown to be equivalent to a point electric dipole moment. The concept of a wavelet will be introduced as the secondary radiation originating from this dipole moment. The collection all such wavelets corresponding to the entire particle volume, will be shown to give the particle's internal and scattered wave; this constitutes the basic premise behind the microphysical model. The large-distance limiting behavior of the wavelet fields will be examined, and from this behavior, the far-field zone approximation will be described. Lastly, a graphical-based method, called phasor analysis, is introduced providing a way to see how the various wavelets of a particle add together to yield the particle's scattered wave.

4.1 Wavelets

Consider what happens to a particle when exposed to the incident plane wave of Eqs. (3.1) and (3.2). The fields of the wave apply forces on the atomic charges of the particle's material via Eq. (1.7). In general, this induces electric and magnetic polarization and current flow.

Comparison of the forms of the Maxwell equations (1.1)-(1.4) with Eqs. (1.22)-(1.25) shows that these effects are described by three[1] distributions of induced sources: a current density $\mathbf{J}^{ind}$, a magnetization current density $\boldsymbol{\nabla} \times \mathbf{M}^{ind}$, and a polarization current density $-i\omega\mathbf{P}^{ind}$ [5, p. 12]. However, because only non-magnetic particles are considered, i.e. $\chi_m = 0$, there is no induced magnetization, and hence no magnetic polarization current density, recall Eq. (1.12). Equation (1.23) can now be expressed as

$$\boldsymbol{\nabla} \times \mathbf{B}^{int}(\mathbf{r}) + i\omega\varepsilon_o\mu_o\mathbf{E}^{int}(\mathbf{r}) = \mu_o\big[\mathbf{J}^{ind}(\mathbf{r}) - i\omega\mathbf{P}^{ind}(\mathbf{r})\big]. \tag{4.1}$$

Using Eqs. (1.10), (1.11), (1.15), (1.22), (1.31), and (1.37), one can reduce Eq. (4.1) to

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}^{int}(\mathbf{r}) - m^2k^2\mathbf{E}^{int}(\mathbf{r}) = 0 \quad \mathbf{r} \in V^{int}, \tag{4.2}$$

which is exactly the same as Eq. (1.36). This shows that the effective sources $\mathbf{J}^{ind}$ and $\mathbf{P}^{ind}$ are collectively described by the internal electric field.

Consider dividing the particle into small volume elements ΔV, each centered on $\mathbf{r}_i$, see Fig. (4.1) below. The size of these elements is small enough that the internal field appears uniform through the element. This division of the particle's volume corresponds to the discretization of the volume integral equations (2.9) and (2.17). Restricting the attention to the internal electric field only[2], the discretization of (2.9) gives

$$\mathbf{E}^{sca}(\mathbf{r}) = \lim_{\Delta V \to 0} k^2(m^2 - 1)\sum_i \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}_i) \cdot \mathbf{E}^{int}(\mathbf{r}_i)\,\Delta V \quad \mathbf{r} \in V^{ext}. \tag{4.3}$$

Comparison of the form of Eq. (4.3) to Eq. (2.29), which leads to Eq. (2.35), shows that each volume element in Eq. (4.3) acts like a point electric dipole. The corresponding expression for the scattered magnetic field follows from Eq. (2.17),

$$\mathbf{B}^{sca}(\mathbf{r}) = -i\omega\mu_o\varepsilon_o(m^2 - 1)\sum_i \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}, \mathbf{r}_i) \cdot \mathbf{E}^{int}(\mathbf{r}_i)\,\Delta V \quad \mathbf{r} \in V^{ext}. \tag{4.4}$$

[1] There can also be an induced polarization charge density $-\boldsymbol{\nabla} \cdot \mathbf{P}^{ind}$, as seen in Eq. (1.24), which will be ignored since it is accounted for by the induced sources $\mu_o\mathbf{J}^{ind}$ and $-i\omega\mathbf{P}^{ind}$ in combination with the continuity relation of Eq. (1.26).

[2] Recall the the magnetic field could also be used here from the curl relationship of Eq. (1.22).

From Eq. (4.3) and the examples presented in Sec. 2.3, one can see that the electric dyadic Green's function $\overleftrightarrow{\mathbf{G}}_e$ performs the role of a propagator that describes how the source within a volume element radiates a field from $\mathbf{r}_i$ to an observation point at $\mathbf{r}$ external to the particle. This source is proportional to the internal field

$$k^2(m^2-1)\mathbf{E}^{int}(\mathbf{r}_i), \tag{4.5}$$

which one can see by comparing Eqs. (4.3) to Eq. (2.8). The wave radiated by this source is called a *wavelet*, and the induced internal-field source of this wavelet, i.e., Eq. (4.5), is the *wavelet source*.

Equation (4.3) and (4.4) are the mathematical embodiment of the *microphysical model*; the response of a particle to the incident wave can be described as a collection of wavelets. Each wavelet is generated by the internal electric field within a differential volume element. The electric and magnetic field of each wavelet are equivalent to that of the wave radiated by a point electric dipole located at the position of the element. The particle's scattered field $\mathbf{E}^{sca}$ is the linear superposition of all of the particle's wavelet fields, and adding this scattered field to the incident gives the total field $\mathbf{E}$ outside of the particle.

Before proceeding, one should note that Eq. (4.3) is valid only outside of the particle volume. The reason for this is because of the singularity that occurs when $\mathbf{r} = \mathbf{r}'$ in the Green's function appearing in the VIE (2.9). Account of this singularity requires one to include an extra dyadic term in Eq. (2.9) and evaluate the principle value of the VIE, see Sec. 5.1. The resulting integral is still a Riemann-type integral and hence can be decomposed into a volume-element sum in exactly the same manner as in Eq. (4.3). The important point here is that this singularity does not invalidate the wavelet superposition concept of the microphysical model.

4.2 Far Field Approximation

Often $\mathbf{r}$ is located in a particle's far-field zone. The mathematical form of the particle's wavelet fields can become substantially simplified depending on just how far from the particle

$\mathbf{r}$ is compared to certain length scales [1, Sec. 3.2], [43]. Consider Eq. (4.3) with the explicit expression for the Green's function of Eq. (2.45) substituted into it,

$$\mathbf{E}^{sca}(\mathbf{r}) = \lim_{\Delta V \to 0} \frac{k^2}{4\pi}(m^2-1) \times \sum_i \frac{\exp(ik\mathfrak{r}_i)}{\mathfrak{r}_i}\left[\left(\frac{1}{k^2\mathfrak{r}_i^2} - \frac{i}{k\mathfrak{r}_i}\right)\left(3\hat{\mathfrak{r}}_i \otimes \hat{\mathfrak{r}}_i - \overleftrightarrow{\mathbf{I}}\right) + \overleftrightarrow{\mathbf{I}} - \hat{\mathfrak{r}}_i \otimes \hat{\mathfrak{r}}_i\right] \cdot \mathbf{E}^{int}(\mathbf{r}_i)\,\Delta V, \quad (4.6)$$

where $\mathfrak{r}_i = \mathbf{r} - \mathbf{r}_i$, see Fig. (4.1). The expression for the magnetic field can be found from Eqs. (2.36) and (4.4) as

$$\mathbf{B}^{sca}(\mathbf{r}) = \lim_{\Delta V \to 0} -i\omega\mu_o\varepsilon_o \frac{k}{4\pi}(m^2-1)\sum_i \frac{\exp(ik\mathfrak{r}_i)}{\mathfrak{r}_i}\left(i - \frac{1}{k\mathfrak{r}_i}\right)\left[\hat{\mathfrak{r}}_i \times \mathbf{E}^{int}(\mathbf{r}_i)\right]\Delta V. \quad (4.7)$$

In addition, suppose that the particle is surrounded by the smallest circumscribing sphere $\mathcal{S}_{sc}$ of radius R_{sc} introduced Sec. 3.1. Then, if $r \gg r_i$,

$$\overleftrightarrow{\mathbf{I}} - \hat{\mathfrak{r}}_i \otimes \hat{\mathfrak{r}}_i \simeq \overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}, \quad (4.8)$$

and since

$$\mathfrak{r}_i = |\mathbf{r} - \mathbf{r}_i| = r\sqrt{1 - 2\frac{\hat{\mathbf{r}} \cdot \mathbf{r}_i}{r} + \frac{r_i^2}{r^2}},$$

one finds that

$$\mathfrak{r}_i \simeq r - \hat{\mathbf{r}} \cdot \mathbf{r}_i + \frac{r_i^2}{2r^2}. \quad (4.9)$$

The next step is to substitute the approximations of Eqs. (4.8) and (4.9) into Eqs. (4.6) and (4.7). However, care must be taken in doing so since certain terms in Eqs. (4.6) and (4.7) have different functional dependencies on $\mathfrak{r}_i$. Since $r \gg r_i$, one can make the substitution of Eq. (4.9) for terms in Eq. (4.6) involving $\mathfrak{r}^{-1}$ and retain only the first term in the expansion of Eq. (4.9), giving

$$\frac{1}{\mathfrak{r}_i} \simeq \frac{1}{r}. \quad (4.10)$$

The exponential term in Eqs. (4.6) and (4.7), however, is more sensitive to $k\mathfrak{r}_i$ and hence the first two terms of Eq. (4.9) are used. Along with Eq. (4.8), Eq. (4.6) becomes

$$\mathbf{E}^{sca}(\mathbf{r}) \simeq \frac{\exp(ikr)}{r}\frac{k^2}{4\pi}(m^2-1)\sum_i \left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}\right) \cdot \mathbf{E}^{int}(\mathbf{r}_i)\exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}_i)\,\Delta V, \quad (4.11)$$

and Eq. (4.7) becomes

$$\mathbf{B}^{sca}(\mathbf{r}) \simeq \omega\mu_o\varepsilon_o \frac{\exp(ikr)}{r} \frac{k}{4\pi}(m^2-1) \sum_i \left[\hat{\mathbf{r}}_i \times \mathbf{E}^{int}(\mathbf{r}_i)\right] \exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}_i)\, \Delta V, \tag{4.12}$$

where the limit appearing in Eqs. (4.3), (4.6), (4.4), and (4.7) is implied and omitted for brevity.

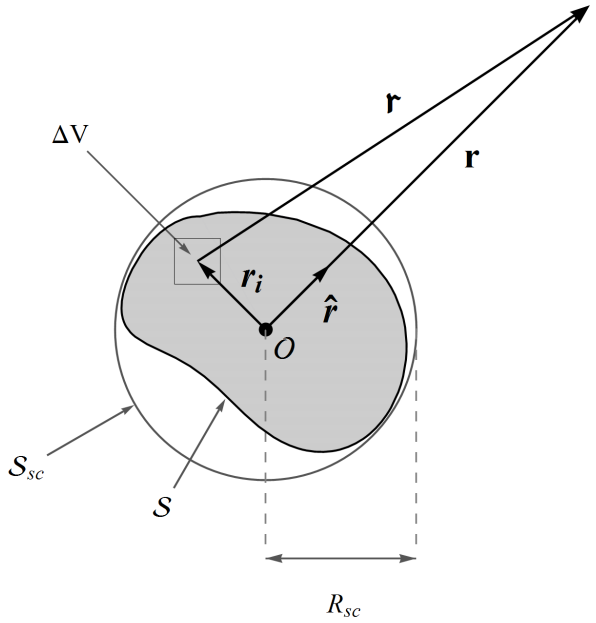

Figure 4.1: Arbitrary particle's wavelet volume element and related vectors and notation.

Equations (4.11) and (4.12) are the *far-field approximation* for the scattered fields in terms of the particle's wavelets. The conditions defining the validity of these approximations are

$$k(r - R_{sc}) \gg 1, \tag{4.13}$$

$$r \gg R_{sc}, \tag{4.14}$$

$$r \gg \frac{kR_{sc}^2}{2}. \tag{4.15}$$

The physical interpretation of these conditions is given in [1, p. 75-77]: Equation (4.13) shows that the distance between the observation point and a wavelet location inside the particle must be much greater than the vacuum wavelength. Equation (4.14) shows that

the distance from the particle to the observation point must be much greater than the overall largest dimension of the particle, given by R_{sc}. Notice that these two conditions, Eq. (4.13) and (4.14), *are not redundant*; the first condition relates r and the particle size to the *wavelength*, whereas the second condition has no dependence on the wavelength. The meaning of the last condition, Eq. (4.15), is explained by considering two wavelet locations $\mathbf{r}_i$ and $\mathbf{r}_j$ and a *fixed* direction to the observation point $\hat{\mathbf{r}}$. The distance between one wavelet and the observation point is $\mathfrak{r}_i$ and the distance between the other wavelet and the observation point is $\mathfrak{r}_j$. Then the meaning of Eq. (4.15) is that the difference in the phase of the wavelets between the path lengths $k\mathfrak{r}_i$ and $k\mathfrak{r}_j$ is independent of the distance r from the particle origin to the observation point, see [1, p. 77] and Fig. (4.2).

An important property of the far-field scattered wave is that it has the form of a spherical wave, recall Sec. 2.3.1. One can see this by defining an angular function $\mathbf{E}_1^{sca}$, called the *scattering amplitude*, from Eq. (4.11) that depends on direction $\hat{\mathbf{r}}$ but is completely independent of distance

$$\mathbf{E}_1^{sca}(\hat{\mathbf{r}}) = \frac{k^2}{4\pi}(m^2-1)\sum_i \left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}_i)\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i)\,\Delta V \quad kr\to\infty. \tag{4.16}$$

Then Eq. (4.11) becomes

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp(ikr)}{r}\mathbf{E}_1^{sca}(\hat{\mathbf{r}}). \tag{4.17}$$

The magnetic field is related to this as

$$\mathbf{B}^{sca}(\mathbf{r}) = \frac{k}{\omega}\frac{\exp(ikr)}{r}\hat{\mathbf{r}}\times\mathbf{E}_1^{sca}(\hat{\mathbf{r}}), \tag{4.18}$$

where Eq. (A.5) has been used. Moreover, using Eq. (A.6), one can see that

$$\hat{\mathbf{r}}\cdot\mathbf{E}^{sca}(\mathbf{r}) = \hat{\mathbf{r}}\cdot\mathbf{B}^{sca}(\mathbf{r}) = 0, \tag{4.19}$$

demonstrating that the far-field scattered wave is transverse. The exponential term $\exp(ikr)/r$ in Eqs. (4.11) and (4.12), along with the tacitly assumed time dependence $\exp(-i\omega t)$, shows that the wave is outward traveling and has surfaces of constant phase that are spheres centered on the particle[3] (and centered on the particle).

[3]Confusion can arise when defining the surfaces of constant phase of the far-field scattered wave. Although

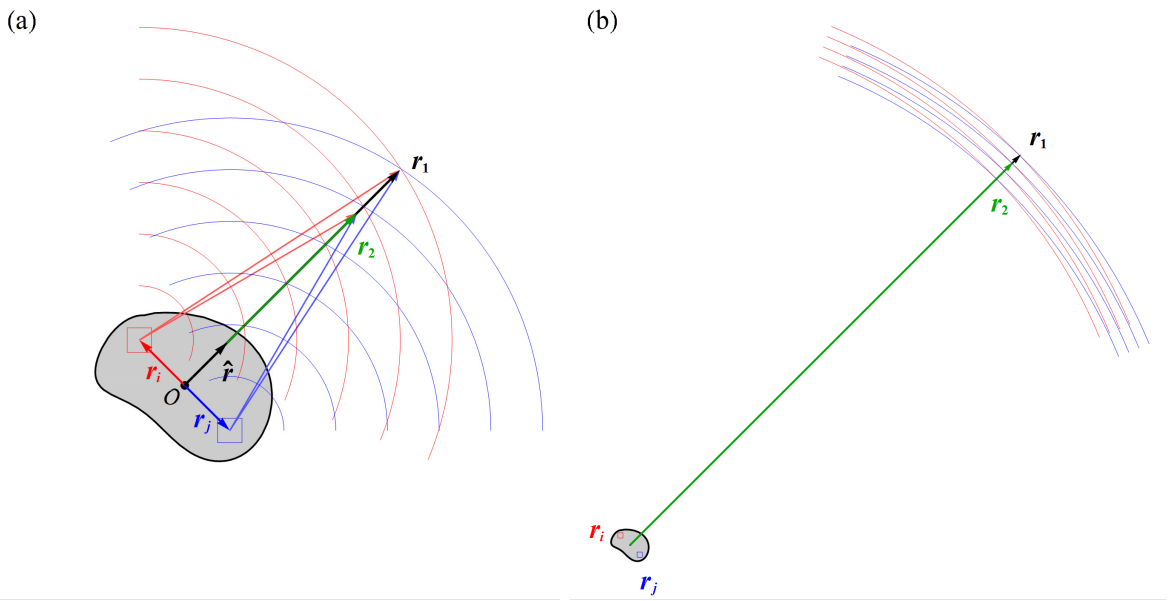

Figure 4.2: Meaning of the far-field approximation based on [1, Sec. 3.2]. In both sketches, the wavefronts are separated by one wavelength intervals and color coded to correspond to their respective locations in the particle.

4.3 Phasors

A unique and important feature of the microphysical model is that the wavelet superposition, which yields the scattered wave, can be presented and analyzed in a visual form. This is achieved by representing the wavelet sums in Eqs. (4.11) and (4.12) as the addition of phasors in the complex plane. To explain how this done in more detail, suppose that the internal field is known and that its components are expressed in the Cartesian basis, i.e.,

$$\mathbf{E}^{int}(\mathbf{r}) = E_x^{int}(\mathbf{r})\,\hat{\mathbf{x}} + E_y^{int}(\mathbf{r})\,\hat{\mathbf{y}} + E_z^{int}(\mathbf{r})\,\hat{\mathbf{z}}.$$

the term $\exp(ikr)/r$ is certainly constant in phase over spherical surfaces, the amplitude function $\mathbf{E}_1^{sca}$ is complex-valued and, in general, varies on these same surfaces. One can then take the surfaces of constant phase to be the spheres and regard $\mathbf{E}_1^{sca}$ as an angular weighting function. Alternatively, one could include the phase of $\mathbf{E}_1^{sca}$ with the term $\exp(ikr)/r$ and hence would not find spherical surfaces of constant phase. It will not matter in this work which tact is taken.

Let the components of the observation point $\mathbf{r}$, however, be represented in the spherical-polar basis with unit vectors

$$\hat{\boldsymbol{r}} = \sin\theta\cos\phi\,\hat{\mathbf{x}} + \sin\theta\sin\phi\,\hat{\mathbf{y}} + \cos\theta\,\hat{\mathbf{z}}, \tag{4.20}$$

$$\hat{\boldsymbol{\theta}} = \cos\theta\cos\phi\,\hat{\mathbf{x}} + \cos\theta\sin\phi\,\hat{\mathbf{y}} - \sin\theta\,\hat{\mathbf{z}}, \tag{4.21}$$

$$\hat{\boldsymbol{\phi}} = -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}}. \tag{4.22}$$

The reason spherical-polar coordinates are chosen here is because the far-field scattered wave is a spherical wave and hence its surfaces of constant phase are degenerate with the surfaces defining the coordinate system. Additionally, many measurement configurations are arranged with spherical geometry around the scattering volume such that, for example, only the polar angle θ dependence of the scattered wave is measured.

Because the far-field wave is transverse, the fields have no $\hat{\mathbf{r}}$ component. Consequently, they can be completely described by their projection onto $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$. To see this, recall that

$$(\hat{\mathbf{r}}\otimes\hat{\mathbf{r}})\cdot\mathbf{E}^{int}(\mathbf{r}_i) = \hat{\mathbf{r}}\left[\hat{\mathbf{r}}\cdot\mathbf{E}^{int}(\mathbf{r}_i)\right]. \tag{4.23}$$

Then,

$$\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}_i) = \mathbf{E}^{int}(\mathbf{r}_i) - \hat{\mathbf{r}}\left[\hat{\mathbf{r}}\cdot\mathbf{E}^{int}(\mathbf{r}_i)\right], \tag{4.24}$$

which from Eqs. (4.20)-(4.22) gives

$$\hat{\mathbf{r}}\cdot\left[\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}_i)\right] = 0, \tag{4.25}$$

$$\hat{\boldsymbol{\theta}}\cdot\left[\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}_i)\right] = \left[\cos\phi\,E_x^{int}(\mathbf{r}_i) + \sin\phi\,E_y^{int}(\mathbf{r}_i)\right]\cos\theta - \sin\theta\,E_z^{int}(\mathbf{r}_i), \tag{4.26}$$

$$\hat{\boldsymbol{\phi}}\cdot\left[\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}_i)\right] = -\sin\phi\,E_x^{int}(\mathbf{r}_i) + \cos\phi\,E_y^{int}(\mathbf{r}_i). \tag{4.27}$$

Now define the far-field wavelet phasors $z_{i\theta}$ and $z_{i\phi}$ as

$$z_{i\theta}(\hat{\mathbf{r}}) = c_o\left\{\left[\cos\phi\,E_x^{int}(\mathbf{r}_i) + \sin\phi\,E_y^{int}(\mathbf{r}_i)\right]\cos\theta - \sin\theta\,E_z^{int}(\mathbf{r}_i)\right\}\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i) \tag{4.28}$$

$$z_{i\phi}(\hat{\mathbf{r}}) = c_o\left[\cos\phi\,E_y^{int}(\mathbf{r}_i) - \sin\phi\,E_x^{int}(\mathbf{r}_i)\right]\exp(-ik\mathbf{r}_i\cdot\hat{\mathbf{r}}), \tag{4.29}$$

where $c_o = k^2(m^2-1)\Delta V/4\pi$. Then Eq. (4.11) becomes

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp ikr}{r} \sum_i \left[z_{i\theta}(\hat{\mathbf{r}})\,\hat{\boldsymbol{\theta}} + z_{i\phi}(\hat{\mathbf{r}})\,\hat{\boldsymbol{\phi}} \right]. \tag{4.30}$$

This shows that in the far-field zone, the particle's scattered wave can be represented as the sum of the complex numbers, $z_{i\theta}$ and $z_{i\phi}$, called phasors. Each of these numbers represents the amplitude of the $\hat{\boldsymbol{\theta}}$ or $\hat{\boldsymbol{\phi}}$ components of the i^{th} wavelet's electric field. These phasors change with direction $\hat{\mathbf{r}}$ but are independent of distance r; the distance dependence is factored out and appears as the exponential coefficient on the sum in Eq. (4.30).

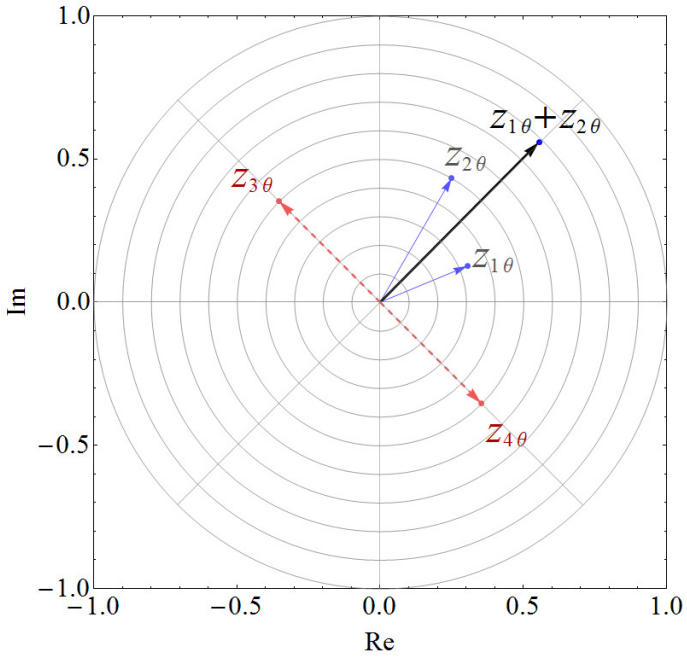

Figure 4.3: Phasor addition in the complex plane.

The primary utility of representing a particle's scattered wave with Eq. (4.30) is that *the phasor addition can be rendered visually to show how the particle's wavelets superimpose to yield the scattered wave*. This is done with phasor plots, where each phasor is a point on the complex plane. With respect to the origin in the complex plane, these points define the tips of vectors that when added yield the component of the net far-field scattered wave[4], see Fig (4.3). In essence, one can look at these phasor plots, one for each the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$

[4]The reader familiar with Huygens' principle may recognize that this phasor analysis is really just a vector-based analog.

components, and see how the phasors add together in the sum in Eq. (4.30). For example, Fig (4.3) shows two pairs of phasors, $(z_{1\theta}, z_{2\theta})$ and $(z_{3\theta}, z_{4\theta})$. Phasors $z_{3\theta}$ and $z_{4\theta}$ are equal in magnitude but shifted by π in phase and hence cancel each other when added together. *This indicates that the wavelets corresponding to these phasors interfere destructively.* On the other hand, phasors $z_{1\theta}$ and $z_{2\theta}$ add together cooperatively, which corresponds to the *constructive interference of the corresponding wavelets.*

In some cases, phasor analysis can reveal how certain aspects of a particles physical character are related to the angular behavior of its scattered wave. This will be the case in Ch. 7 where it will be shown that phasor analysis explains the meaning of the transitions between different angular power-law-like regions of a particle's scattered intensity curve. As a final note, although the phasor analysis described above applies only to the far-field scattered wave, the concept is certainly not restricted to the far-field zone. One can easily generalize this analysis method to apply to the near-field wave by defining the phasors from Eqs. (4.6) and (4.7), rather than from Eq. (4.16). Notice though, that this will now require one to consider a third phasor to account for the possibility that the fields have a radial component.

Chapter 5

Discrete Dipole Approximation

"The mathematics of the full treatment may be altogether beyond human power in a reasonable time; nevertheless..."

-O. Heaviside [49, p. 324]

This chapter shows how the VIE can be solved for the internal field of an arbitrarily-shaped particle. The solution is achieved by discretizing the VIE over the particle volume in terms of a cubic lattice of electric dipole moments; this is the essence of the so-called Discrete Dipole Approximation (DDA). The discretization leads to a linear system of coupled algebraic equations for the moments, which are then solved using an iterative numerical procedure. In the process of developing the DDA, one is introduced to the concept of coupling, which leads to an enhanced physical picture of the microphysical model. For example, one will see that coupling is the cause of refraction, and as such, one is reminded that refraction *does not* correspond to the slowing of the speed of light inside a medium.

5.1 Volume Integral Equation Revisited

The theoretical foundation of the microphysical model presented in Ch. 4 is the volume integral equation of (2.10);

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + k^2(m^2-1)\int_{V^{int}} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}^{int}(\mathbf{r}')\,\mathrm{d}\mathbf{r}' \quad \mathbf{r}\in V^{ext}\cup V^{int},$$

where the electric dyadic Green's function $\overleftrightarrow{\mathbf{G}}_e$, given by Eq. (2.45), is

$$\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') = \frac{\exp(ik\mathfrak{r})}{4\pi\mathfrak{r}}\left[\left(\frac{1}{k^2\mathfrak{r}^2}-\frac{i}{k\mathfrak{r}}\right)\left(3\hat{\mathfrak{r}}\otimes\hat{\mathfrak{r}}-\overleftrightarrow{\mathbf{I}}\right)+\overleftrightarrow{\mathbf{I}}-\hat{\mathfrak{r}}\otimes\hat{\mathfrak{r}}\right],$$

and

$$\mathfrak{r} = \mathbf{r}-\mathbf{r}',$$

is the vector between the observation point $\mathbf{r}$ and the integration point $\mathbf{r}'$ within the particle volume V^{int}. Discretization of this volume integral equation leads to the description of a particle's far-field scattered wave in Eq. (4.11) in terms wavelet superposition. These wavelets, however, are determined by the internal field which is unknown at this point.

One way to solve the VIE (2.10) for the internal field is to follow the same discretization procedure as is done in Sec. 4.1, except now the observation point is placed inside the particle. Doing this causes problems for the integral equation. When $\mathbf{r}\in V^{int}$, there will eventually be a point in the evaluation of the volume integral where $\mathbf{r}=\mathbf{r}'$. This causes the Green function to diverge with first, second, and third order singularities, as one can see from the powers on $\mathfrak{r}$ in Eq. (2.45). Following [50, Sec. 1.3] and [51], the singularity in VIE is accounted for by the exclusion of a small volume V_δ enclosing the singular point at $\mathbf{r}=\mathbf{r}'$. This exclusion requires the addition of two dyadic terms to the remaining integral. The first term, $\overleftrightarrow{\mathbf{M}}$, accounts for the effect on the field at $\mathbf{r}=\mathbf{r}'$, which is due to the finite size of the excluded volume V_δ. The second term, $\overleftrightarrow{\mathbf{L}}$, is the self-interaction dyadic[1], which accounts for the electromagnetic interaction of the excluded volume with itself, see below. The VIE then becomes

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r})+k^2(m^2-1)\int_{V^{int}-V_\delta}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}^{int}(\mathbf{r}')\,\mathrm{d}\mathbf{r}'+\left(\overleftrightarrow{\mathbf{M}}-\overleftrightarrow{\mathbf{L}}\right)\cdot\mathbf{E}(\mathbf{r})\quad \mathbf{r}\in\cup V^{int}. \tag{5.1}$$

Note that the observation point is now restricted to be inside the particle. Both $\overleftrightarrow{\mathbf{M}}$ and $\overleftrightarrow{\mathbf{L}}$ are independent of location, but depend on the shape of the excluded volume V_δ. If

[1]This term is also called the "self-term," and the "depolarizing dyadic," see [51] and [6, p. 370], respectively.

V_δ is taken to be one of the cubic volume elements ΔV described below, then $\overleftrightarrow{\mathbf{M}}$ can be approximately evaluated [30, Sec. 3.21], [51]

$$\overleftrightarrow{\mathbf{M}} \simeq \left[\left(\frac{4\pi}{3} \right)^{\frac{1}{3}} (kd)^2 + \frac{2i}{3} (kd)^3 \right] \overleftrightarrow{\mathbf{I}}, \tag{5.2}$$

and

$$\overleftrightarrow{\mathbf{L}} = \frac{4\pi}{3} \overleftrightarrow{\mathbf{I}}. \tag{5.3}$$

Now take the division of the particle in Sec. 4.1 to be done such that the volume elements ΔV are cubes of side d. The centers of the elements form a cubical lattice of spacing d where the i^{th} site is denoted by $\mathbf{r}_i$. Then, with Eq. (5.3), Eq. (5.1) becomes

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + k^2(m^2 - 1) \sum_i^N \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}_i) \cdot \mathbf{E}^{int}(\mathbf{r}_i) \, \Delta V + \left(\overleftrightarrow{\mathbf{M}} - \overleftrightarrow{\mathbf{L}} \right) \cdot \mathbf{E}(\mathbf{r}), \quad \mathbf{r} \neq \mathbf{r}_i, \tag{5.4}$$

where N is the total number volume elements included in V^{int}.

The field $\mathbf{E}^{int}$ in Eq. (5.4) is still unknown. However, a system of coupled linear algebraic equations can be formed from Eq. (5.4) by placing the observation point $\mathbf{r}$ on each lattice site. A convenient way to express this linear system is to cast the equations in terms of the electric dipole moment corresponding to the polarization of each volume element. Recall from Sec. 4.1 that the induced current and polarization are related to the internal electric field. If the element is small enough that the internal field is suitably uniform throughout, then the induced sources within the element can be represented by a single electric dipole moment $\mathbf{p}_i$. This moment is related to the internal electric field as

$$\mathbf{p}_i = \alpha \mathbf{E}^{int}(\mathbf{r}_i), \tag{5.5}$$

where α is the scalar[2] electric polarizability.

Now consider a new dyadic $\overleftrightarrow{\mathbf{A}}$, which is proportional to the electric dyadic Green's

[2]The more general situation would correspond to a birefringent, anisotropic, and inhomogeneous particle composition. In that case the polarizability would be a 3×3 dyadic with nine independent elements.

function of Eq. (2.45) as

$$\begin{aligned}\overleftrightarrow{\mathbf{A}}(\mathbf{r}_i,\mathbf{r}_j) &= -\frac{k^2}{\varepsilon_o}\overleftrightarrow{\mathbf{G}}_e(\mathbf{r}_i,\mathbf{r}_j)\\ &= \frac{\exp(ik\mathfrak{r}_{ij})}{4\pi\varepsilon_o\mathfrak{r}_{ij}}\left[k^2\left(\hat{\mathfrak{r}}_{ij}\otimes\hat{\mathfrak{r}}_{ij}-\overleftrightarrow{\mathbf{I}}\right)+\frac{1-ik\mathfrak{r}_{ij}}{\mathfrak{r}_{ij}^2}\left(\overleftrightarrow{\mathbf{I}}-3\hat{\mathfrak{r}}_{ij}\otimes\hat{\mathfrak{r}}_{ij}\right)\right],\end{aligned} \tag{5.6}$$

where

$$\mathfrak{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j \tag{5.7}$$

is the difference vector between the i^{th} and j^{th} lattice site. With Eqs. (5.5)-(5.7), Eq. (5.4) becomes

$$\mathbf{p}_i = \alpha\mathbf{E}^{inc}(\mathbf{r}_i) - \alpha\sum_{\substack{j=1\\ i\neq j}}^{N}\overleftrightarrow{\mathbf{A}}(\mathbf{r}_i,\mathbf{r}_j)\cdot\mathbf{p}_j\,, \tag{5.8}$$

see [50, p. 27-31] for more detail. Using the Kronecker delta function δ_{ij},

$$\delta_{ij} = \begin{cases} 1 & i=j\\ 0 & i\neq j,\end{cases} \tag{5.9}$$

Eq. (5.8) can be expressed in the more compact form

$$\mathbf{E}^{inc}(\mathbf{r}_i) = \sum_{j=1}^{N}\left[\frac{1}{\alpha}\delta_{ij}\overleftrightarrow{\mathbf{I}}+(1-\delta_{ij})\overleftrightarrow{\mathbf{A}}(\mathbf{r}_i,\mathbf{r}_j)\right]\cdot\mathbf{p}_j. \tag{5.10}$$

The discretization-based solution of the VIE given by this system of coupled linear *algebraic* equations is known as the Discrete Dipole Approximation[3] (DDA). An efficient numerical method to solve this linear system of equations will be presented in Sec. 5.2 and the physical significance of the DDA system of equations will be described in Sec. 5.6. Once the dipole moments are known, the internal electric field follows from Eq. (5.5) and the scattered field can be calculated from the moments as discussed in Sec. 5.3.

The singularity from the Greens function in the VIE is accounted for by the relation between α and the refractive index m. In the static limit, $\omega\to 0$, this relationship is given by the Clausius-Mosotti (CM) expression [6, p. 217]

$$\alpha_o = 3\varepsilon_o d^3\left(\frac{m^2-1}{m^2+2}\right). \tag{5.11}$$

[3]The DDA was originally formulated by [52] *a priori*. It was not until more recently that it was formally derived from the VIE [51].

It is important to emphasize that Eq. (5.11) is valid only in the static limit and hence cannot apply in Eqs. (5.5) and (5.10). When the frequency is nonzero, the wavelength is not infinite and there is inherent error in assuming that the field throughout ΔV is uniform. Consequently, α must be corrected for the finiteness of the element's size, and this correction is accounted for by the introduction of the $\overleftrightarrow{\mathbf{M}}$ dyadic in Eq. (5.4).

There is a difficult problem associated with the self-force, or radiation reaction, of a point dipole. Radiation reaction becomes important when one tries to determine *how* an oscillating electric field induces a volume element's dipole moment. The affect of radiation reaction is that the dipole "feels" a resistance to its oscillatory motion, and this results in *both* damping and a shift in phase between the oscillation and the driving field[4]. One will see in Chs. 10 and 11 that this phase shift is required in order for the dipole to redistribute energy from the wave driving it into the secondary wave, i.e. its wavelet. Then, in non-static situations *the polarizability must be a complex-valued quantity* in order for it to be able to describe the damping and phase shift caused by radiation reaction. This is true even for completely real refractive indices.

Corrections to the CM polarizability that account for the finiteness of the volume elements and the affect of radiation reaction are discussed in [51] and are ultimately related to the addition of the $\overleftrightarrow{\mathbf{M}}$ and $\overleftrightarrow{\mathbf{L}}$ dyadics. The correction used throughout this work is derived in [50, Sec. 3.2] and gives

$$\alpha = \frac{\alpha_o}{1 - \frac{\alpha_o}{4\pi\varepsilon_o d^3}\left[\left(\frac{4\pi}{3}\right)^{\frac{1}{3}} k^2 d^2 + \frac{2ik^3 d^3}{3}\right]}. \tag{5.12}$$

Using the notation of [51], this α will be called the Digitized Green's Function (DGF) polarizability. Another commonly used correction for the polarizability is the Lattice Dispersion Relation (LDR), which has the same order of correction as the DGF polarizability, i.e., $(kd)^3$, but also takes into account the geometry of the discretization lattice with respect to the incident wave. A study showing the multiple formulations for α and their resulting

[4]The physical origin of radiation reaction remains one of the enduring mysteries in theoretical physics [53, p. 32-1].

errors, as compared to Mie theory, is given in [51]. In general the LDR is more accurate than the DGF, but the difference is small enough (a few percent error) that only the DGF polarizability is used here.

5.2 Solving the System of Equations

The N coupled linear algebraic equations given by Eq. (5.10) can be cast into the form of a matrix equation. The advantage of doing so is that standard methods can be used to numerically obtain the solutions for the dipole moments. Moreover, the discretization of the VIE on a cubic lattice allows one to solve the system using a Fast Fourier Transform (FFT) based iterative procedure, which is much faster than direct inversion of the matrix equation. The following will briefly describe how this is done, while the numerical methods of the solution are provided in App. C and D

To begin, let $\bar{\mathbf{E}}^{inc}$ represent a column vector with each element given by the value of the incident electric field at a lattice site,

$$\bar{\mathbf{E}}^{inc} = \begin{pmatrix} \mathbf{E}^{inc}(\mathbf{r}_1) \\ \mathbf{E}^{inc}(\mathbf{r}_2) \\ \mathbf{E}^{inc}(\mathbf{r}_3) \\ \vdots \\ \mathbf{E}^{inc}(\mathbf{r}_N) \end{pmatrix}. \tag{5.13}$$

Similarly, let $\bar{\mathbf{p}}$ represent a column vector with elements given by the (currently unknown) dipole moments

$$\bar{\mathbf{p}} = \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \\ \vdots \\ \mathbf{p}_N \end{pmatrix}. \tag{5.14}$$

Notice that each element of these column vectors is itself a column vector consisting of the

three Cartesian components of the vector, i.e.,

$$\bar{\mathbf{p}} = \begin{pmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \\ \vdots \\ \mathbf{p}_N \end{pmatrix} = \begin{pmatrix} p_{x1} \\ p_{y1} \\ p_{z1} \\ p_{x2} \\ p_{y2} \\ p_{z2} \\ \vdots \\ p_{xN} \\ p_{yN} \\ p_{zN} \end{pmatrix}.$$

The dimension of the column vectors in Eqs. (5.13) and (5.14) is $3N$. Returning to Eq. (5.10), one can now see that the column vectors of Eqs. (5.13) and (5.14) are part of the matrix equation

$$\bar{\mathbf{E}}^{inc} = \bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}, \tag{5.15}$$

where the elements of the coefficient matrix $\bar{\bar{\mathbf{A}}}$ are the dyadic coefficients appearing in Eq. (5.10),

$$\left[\bar{\bar{\mathbf{A}}}\right]_{ij} = \frac{1}{\alpha}\delta_{ij}\overleftrightarrow{\mathbf{I}} + (1-\delta_{ij})\overleftrightarrow{\mathbf{A}}(\mathbf{r}_i, \mathbf{r}_j), \tag{5.16}$$

or, more explicitly in matrix form,

$$\bar{\bar{\mathbf{A}}} = \begin{pmatrix} \frac{1}{\alpha}\overleftrightarrow{\mathbf{I}} & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_1, \mathbf{r}_2) & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_1, \mathbf{r}_3) & \dots & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_1, \mathbf{r}_N) \\ \overleftrightarrow{\mathbf{A}}(\mathbf{r}_2, \mathbf{r}_1) & \frac{1}{\alpha}\overleftrightarrow{\mathbf{I}} & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_2, \mathbf{r}_3) & \dots & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_2, \mathbf{r}_N) \\ \overleftrightarrow{\mathbf{A}}(\mathbf{r}_3, \mathbf{r}_1) & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_3, \mathbf{r}_2) & \frac{1}{\alpha}\overleftrightarrow{\mathbf{I}} & \dots & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_3, \mathbf{r}_N) \\ \overleftrightarrow{\mathbf{A}}(\mathbf{r}_N, \mathbf{r}_1) & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_N, \mathbf{r}_2) & \overleftrightarrow{\mathbf{A}}(\mathbf{r}_N, \mathbf{r}_3) & \dots & \frac{1}{\alpha}\overleftrightarrow{\mathbf{I}} \end{pmatrix}.$$

One can see from Eq. (5.16) that the number of equations is $3N$ and that the dimension of the coefficient matrix $\bar{\bar{\mathbf{A}}}$ is $3N \times 3N$. In the course of deriving Eq. (5.16), it is assumed that the size of a volume element is small enough that the internal field inside it is uniform and constant. The validity of this assumption will rely on how finely the particle volume is discretized with respect to the overall variation of the actual internal field throughout its interior, and with respect to how well the resulting discretization lattice mimics the particle surface. Therefore, it is the fineness of the discretization that primarily determines

the DDA's accuracy. The problem is that the size of the matrix system increases rapidly as the discretization is made finer, i.e., as N increases. The typical number of volume elements required to accurately solve for the scattered fields for a several-micron sized particle illuminated by 550 nm light can be as large as ten million. This corresponds to a dense coefficient matrix with 9×10^{14} elements! Such a matrix would require tera-bites of memory and cannot currently be inverted by any computer on human time scales. Fortunately, the matrix has a high degree of symmetry, which will enable use of a fast numerical iterative method to solve Eq. (5.15) without having to calculate and store the majority of coefficient matrix elements.

The first crucial observation of the structure of $\bar{\bar{\mathbf{A}}}$ is that it is symmetric about its main diagonal. To see this symmetry, recall that the functional form of the off diagonal elements are given by the dyadic $\overleftrightarrow{\mathbf{A}}$,

$$\overleftrightarrow{\mathbf{A}}(\mathbf{r}_i, \mathbf{r}_j) = \frac{\exp(ik\mathfrak{r}_{ij})}{4\pi\varepsilon_o \mathfrak{r}_{ij}} \left[k^2 \left(\hat{\mathfrak{r}}_{ij} \otimes \hat{\mathfrak{r}}_{ij} - \overleftrightarrow{\mathbf{I}} \right) + \frac{1 - ik\mathfrak{r}_{ij}}{\mathfrak{r}_{ij}^2} \left(\overleftrightarrow{\mathbf{I}} - 3\hat{\mathfrak{r}}_{ij} \otimes \hat{\mathfrak{r}}_{ij} \right) \right].$$

The transpose of $\bar{\bar{\mathbf{A}}}$ is facilitated by reversing the order of the indices, i.e. $ij \rightarrow ji$ in Eq. (5.16). This operation is equivalent to making the substitution $\mathbf{r}_i \rightarrow -\mathbf{r}_i$ and $\mathbf{r}_j \rightarrow -\mathbf{r}_j$ and one can see that the elements are unaffected by this substitution since $\mathfrak{r}_{ij}$ occurs in $\bar{\bar{\mathbf{A}}}$ only as its magnitude or as a dyadic product with itself. Consequently, $\bar{\bar{\mathbf{A}}}^{\mathrm{T}} = \bar{\bar{\mathbf{A}}}$, where $\bar{\bar{\mathbf{A}}}^{\mathrm{T}}$ denotes the transpose of the matrix. Additionally, one can see that this also means that

$$\bar{\bar{\mathbb{A}}}^{\dagger} \cdot \mathbf{p}_i = \left[\bar{\bar{\mathbb{A}}} \cdot \mathbf{p}_i^* \right]^*, \tag{5.17}$$

where $\dagger$ denotes the Hermitian conjugate. The usefulness of Eq. (5.17) is that the Hermitian conjugate does not have to be explicitly calculated in the numerical iterative solution of Eq. (5.15), and is discussed in App. C.

The next important property of $\bar{\bar{\mathbf{A}}}$ is the translational invariance of its constituent dyadic Green's function $\overleftrightarrow{\mathbf{G}}_e$. The translational invariance is demonstrated by shifting the coordinate origin by an arbitrary vector $\mathbf{a}$. Doing so results in shifting $\mathbf{r}_i$ and $\mathbf{r}_j$ by $\mathbf{r}_i \rightarrow \mathbf{r}_i - \mathbf{a}$ and $\mathbf{r}_j \rightarrow \mathbf{r}_j - \mathbf{a}$. Referring to Eq. (5.6), one can see that this operation leaves the matrix

elements unaffected since $\mathbf{a}$ cancels out in both r_{ij} and $\hat{\mathbf{t}}_{ij} \otimes \hat{\mathbf{t}}_{ij}$. The consequence of the translational invariance of $\bar{\bar{\mathbf{A}}}$ means that the matrix elements depend only on the *relative* separation between the lattice sites.

The location of a lattice site can be described by three integer indices representing the Cartesian components of the site's location. For example, let the location of the i^{th} site be

$$\mathbf{r}_i = n\,\hat{\mathbf{x}} + m\,\hat{\mathbf{y}} + l\,\hat{\mathbf{z}}, \quad n, m, l \in \{1, 2, 3, \ldots \mathfrak{n}\},$$

and the j^{th} site,

$$\mathbf{r}_j = n'\,\hat{\mathbf{x}} + m'\,\hat{\mathbf{y}} + l'\,\hat{\mathbf{z}}, \quad n', m', l' \in \{1, 2, 3, \ldots \mathfrak{n}\},$$

where $\mathfrak{n}$ is the size of the lattice, i.e. $N = \mathfrak{n} \times \mathfrak{n} \times \mathfrak{n}$. Then, the translational invariance of the coefficient matrix means that

$$\bar{\bar{\mathbf{A}}}_{ij} = \bar{\bar{\mathbf{A}}}_{nn'mm'll'} = \bar{\bar{\mathbf{A}}}_{n-n',m-m',l-l'}. \tag{5.18}$$

A matrix with the property of Eq. (5.18) is known as a Toeplitz matrix [50]. Rewriting Eq. (5.15) in summation form

$$\mathbf{E}^{inc}(n, m, l) = \sum_{n',m',l'=1}^{\mathfrak{n}} \left[\bar{\bar{\mathbf{A}}}\right]_{n-n',m-m',l-l'} \cdot [\mathbf{p}]_{n'm'l'}\,, \tag{5.19}$$

shows that the matrix-vector product $\bar{\bar{\mathbb{A}}} \cdot \bar{\mathbf{p}}$ is equivalent to a discrete convolution in three dimensions. This is a substantial computational advantage because discrete convolutions can be evaluated with fast Fourier transforms (FFT), which involve $\sim N \log N$ operations as opposed to $(3N)^2$ for the direct multiplication. Appendices C and D demonstrate how one can use FFTs in combination with an iterative minimization procedure to numerically solve Eq. (5.19).

5.3 Near and Far-Field Scattered Wave

Once the complex-valued Cartesian components of the dipole moments have been found from Eq. (5.19), the internal electric field can be found from the moments via Eq. (5.5).

From the internal field, one can calculate a number of scattering quantities, e.g., the far-field phase function[5], total scattering, absorption, and extinction cross sections, and Stokes parameters. Provided that the observation point is outside of the particle, (as it is for the calculation of these scattering quantities), the scattered field is given in terms of the dipole moments as

$$\begin{aligned}\mathbf{E}^{sca}(\mathbf{r}) &= -\sum_i^N \overleftrightarrow{\mathbf{A}}(\mathbf{r},\mathbf{r}_i)\cdot\mathbf{p}_i \\ &= \sum_i^N \frac{\exp(ik\mathfrak{r}_i)}{4\pi\varepsilon_o\mathfrak{r}_i}\left[k^2\left(\overleftrightarrow{\mathbf{I}}-\hat{\mathfrak{r}}_i\otimes\hat{\mathfrak{r}}_i\right)+\frac{1-ik\mathfrak{r}_i}{\mathfrak{r}_i^2}\left(3\hat{\mathfrak{r}}_i\otimes\hat{\mathfrak{r}}_i-\overleftrightarrow{\mathbf{I}}\right)\right]\cdot\mathbf{p}_i \quad \mathbf{r}\in V^{ext}, \quad (5.20)\end{aligned}$$

where $\mathfrak{r}_i = \mathbf{r}-\mathbf{r}_i$. Notice that this expression is valid at any distance from the particle surface; the far-field approximation has not yet been imposed. The scattered magnetic field follows from Eqs. (2.36), (4.4), and (5.8),

$$\begin{aligned}\mathbf{B}^{sca}(\mathbf{r}) &= -i\omega\mu_o\sum_i^N \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}_i)\cdot\mathbf{p}_i \\ &= -i\omega\mu_o\sum_i^N \frac{\exp(ik\mathfrak{r}_i)}{4\pi\mathfrak{r}_i}\left(ik-\frac{1}{\mathfrak{r}_i}\right)\hat{\mathfrak{r}}_i\times\mathbf{p}_i \quad \mathbf{r}\in V^{ext}, \quad (5.21)\end{aligned}$$

and is also valid at any distance from the particle.

In the far-field zone, Eq. (5.20) simplifies to

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{k^2\exp(ikr)}{4\pi\varepsilon_o r}\sum_i^N\left(\overleftrightarrow{\mathbf{I}}-\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{p}_i\,\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i). \quad (5.22)$$

The scattering amplitude can then be identified from Eq. (5.22) by analogy to Eq. (4.17),

$$\mathbf{E}_1^{sca}(\hat{\mathbf{r}}) = \frac{k^2}{4\pi\varepsilon_o}\sum_i^N\left(\overleftrightarrow{\mathbf{I}}-\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{p}_i\,\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i), \quad (5.23)$$

such that

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp(ikr)}{r}\mathbf{E}_1^{sca}(\hat{\mathbf{r}}).$$

From Eq. (4.18), the corresponding magnetic field is

$$\mathbf{B}^{sca}(\mathbf{r}) = \frac{\omega\mu_o k}{4\pi}\frac{\exp(ikr)}{r}\sum_i^N(\hat{\mathbf{r}}\times\mathbf{p}_i)\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i). \quad (5.24)$$

[5]Also called the scattered intensity.

5.4 Orientational Averaging

It is often of interest to calculate the orientational average of a particle's scattering quantity as it assumes some distribution of orientations. For example, Ch. 8 will study an ensemble of identical, randomly oriented, non-interacting fractal-like aggregate particles. The scattering properties of such a system can be approximated by the statistical ensemble average of a single aggregate's orientation [1, Secs. 5.1-5.3], [54, p. 34]. However, both the internal and scattered fields for a such a particle depend on its orientation with respect to the incident wave. Consequently, the dipole moments in the DDA will be different for each orientation. The following will demonstrate how the orientational ensemble average is described and calculated with the DDA.

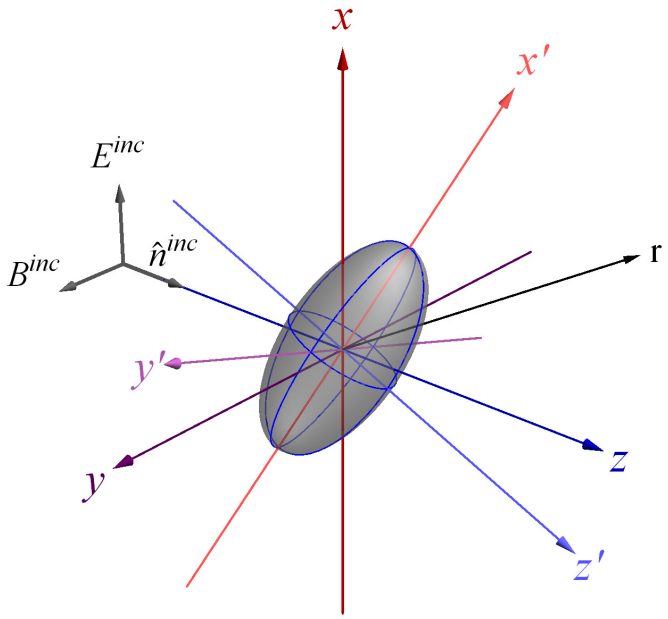

Figure 5.1: Laboratory coordinate system (LCS), indicated by x, y, and z, and the particle coordinate system (PCS), indicated by x', y', and z'.

The first task is to describe the orientation of the particle with respect to the incident wave. This is achieved by introducing two coordinate systems; the Laboratory Coordinate System (LCS) and the Particle Coordinate System (PCS). The LCS is fixed to the prop-

agation and polarization direction of the incident wave whereas the PCS is fixed to the particle, see Fig. (5.1). The orientation of one coordinate systems with respect to the other is described with the Euler angles,

$$\alpha \in [0, 2\pi), \quad \beta \in [0, \pi], \quad \gamma \in [0, 2\pi). \tag{5.25}$$

See [2, Sec. 2.4] for more details. Let $\hat{\mathbf{r}}_L$ and $\hat{\mathbf{r}}_P$ denote the same direction as described in the LCS and PCS, respectively. Additionally, let (θ_L, ϕ_L) and (θ_P, ϕ_P) be the polar and azimuthal angles associated with $\hat{\mathbf{r}}_L$ and $\hat{\mathbf{r}}_P$;

$$\hat{\mathbf{r}}_L = \begin{pmatrix} \sin\theta_L \cos\phi_L \\ \sin\theta_L \sin\phi_L \\ \cos\theta_L \end{pmatrix}, \quad \hat{\mathbf{r}}_P = \begin{pmatrix} \sin\theta_P \cos\phi_P \\ \sin\theta_P \sin\phi_P \\ \cos\theta_P \end{pmatrix}. \tag{5.26}$$

The two vectors are then related by the Euler rotation matrix $\overleftrightarrow{\boldsymbol{\beta}}$

$$\overleftrightarrow{\boldsymbol{\beta}}(\alpha, \beta, \gamma) = \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\beta\cos\gamma \\ -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\beta\sin\gamma \\ \cos\alpha\sin\beta & \sin\alpha\sin\beta & \cos\beta \end{pmatrix} \tag{5.27}$$

as

$$\hat{\mathbf{r}}_P = \overleftrightarrow{\boldsymbol{\beta}}(\alpha, \beta, \gamma) \cdot \hat{\mathbf{r}}_L \tag{5.28}$$

$$\hat{\mathbf{r}}_L = \left[\overleftrightarrow{\boldsymbol{\beta}}(\alpha, \beta, \gamma)\right]^{-1} \cdot \hat{\mathbf{r}}_P, \tag{5.29}$$

where $\left[\overleftrightarrow{\boldsymbol{\beta}}\right]^{-1}$ is the inverse of Eq. (5.27),

$$\left[\overleftrightarrow{\boldsymbol{\beta}}(\alpha, \beta, \gamma)\right]^{-1} = \begin{pmatrix} \cos\alpha\cos\beta\cos\gamma - \sin\alpha\sin\gamma & -\cos\alpha\cos\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta \\ \sin\alpha\cos\beta\cos\gamma + \cos\alpha\sin\gamma & -\sin\alpha\cos\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta \\ -\sin\beta\cos\gamma & \sin\beta\sin\gamma & \cos\beta \end{pmatrix}. \tag{5.30}$$

Notice that $\left[\overleftrightarrow{\boldsymbol{\beta}}\right]^{-1} = \left[\overleftrightarrow{\boldsymbol{\beta}}\right]^{T}$.

A change in particle orientation requires that it be rotated from its original position in one of two ways: It can be rotated relative to the dipole lattice, in which case the lattice

is fixed to the LCS. Alternatively, the lattice can be fixed to the particle, and then the incident wave propagation and polarization directions can be rotated. This latter method is preferable since the shape errors caused by the limited fineness of the lattice do not change with orientation, see Sec. 5.5 below. A subtle consequence of this latter rotation method is that the dipole moments must then be rotated into the LCS before any scattering quantities can be calculated from them.

Now consider an arbitrary quantity Q that depends on particle orientation as described by (α, β, γ). For example, Q could be one of the Stokes parameters, or one of the total cross sections. Then, the orientational average of this quantity over an infinite set of equally probable[6] random particle-orientations is given by the ensemble average [2]

$$\langle Q(\hat{\mathbf{r}})\rangle_{ori} = \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} Q(\hat{\mathbf{r}};\alpha,\beta,\gamma)\sin\beta\,\mathrm{d}\alpha\,\mathrm{d}\beta\,\mathrm{d}\gamma, \tag{5.31}$$

where the factor $1/8\pi^2$ accounts for normalization.

If the functional dependence of Q on the Euler angles is known, then it is possible that Eq. (5.31) can be evaluated analytically, resulting in an exact ensemble average. This can be done using the T-Matrix method but cannot currently be done in the DDA since the solutions for the dipole moments are known only numerically for a *single* orientation. Consequently, Eq. (5.31) must be discretized and evaluated by performing the numerical solution for each orientation of the discretized set. This is one of the primary disadvantages to the DDA; the discretized averaging does not yield an exact result and can be very time consuming depending on the size, refractive index, and complexity of the particle shape.

Work by [55] shows that DDA-based analytical orientational averaging is possible *if* one is able to find the inverse of the coefficient matrix in Eq. (5.15). However, this is not currently very useful since numerical matrix inversion methods scale in computational time as the square of the matrix dimension, and that can be excessively large even for moderately sized particles. Another way to build-in orientational averaging in the DDA is to use the

[6] An orientational probability distribution function can be included in the integral to describe a distribution with a preferred direction, e.g., as would be the case for gravitational alignment of the particles.

DDA to solve for the coefficients of the T-Matrix. This then allows one to exploit the same analytical averaging methods to calculate the ensemble average [56]. Although this method is very clever and probably the only practical way to do analytical averaging in the DDA, it obscures the connection between the internal field (dipoles) and the corresponding scattering quantity. Consequently, this method is not pursued here.

To describe how the discretized orientational averaging is done in this work, consider N^{ori} Euler angles sampled in the intervals given in Eq. (5.25). This corresponds to N_{ori} particle orientations with the i^{th} orientation described by $(\alpha_i, \beta_i, \gamma_i)$. Then Eq. (5.31) becomes

$$\langle Q(\hat{\mathbf{r}})\rangle_{ori} \simeq \frac{1}{8\pi^2} \sum_{i}^{N_{ori}} Q(\hat{\mathbf{r}})_i \sin\beta_i \, \Delta\alpha \, \Delta d\beta \, \Delta d\gamma. \tag{5.32}$$

Now one must decide how to select the discrete orientations to correspond to a uniform distribution. One possibility would be to sample the Euler angles at random uniformly N^{ori} times within each angle's respective interval. However, this sampling method results in a non-uniform distribution of particle orientations, see (a) in Fig. (5.2). A proper sampling of the Euler angles corresponding to a uniform random distribution is given by [57, p. 91] and consists of a non-uniform sampling of the Euler angles described by

$$\alpha_i = 2\pi x_1, \quad \beta_i = \cos^{-1}(2x_2 - 1), \quad \gamma_i = 2\pi x_3, \tag{5.33}$$

where (x_1, x_2, x_3) are random numbers uniformly sampled in $[0, 1]$, see plot (b) in Fig. (5.2). Another reason that the discrete orientations should be chosen by sampling the Euler angles randomly via Eq. (5.33) is to avoid beating effects that can result from the angular dependence of Q with the mesh size associated with a sampling of the Euler angles at constant intervals.

Another issue associated with orientational averaging in the DDA concerns the number of orientations that must be considered before the averaged quantity converges to its true analytical value. Because the numerical solution for the dipole moments can be quite time consuming, it is advantageous to compute the smallest number of particle orientations as

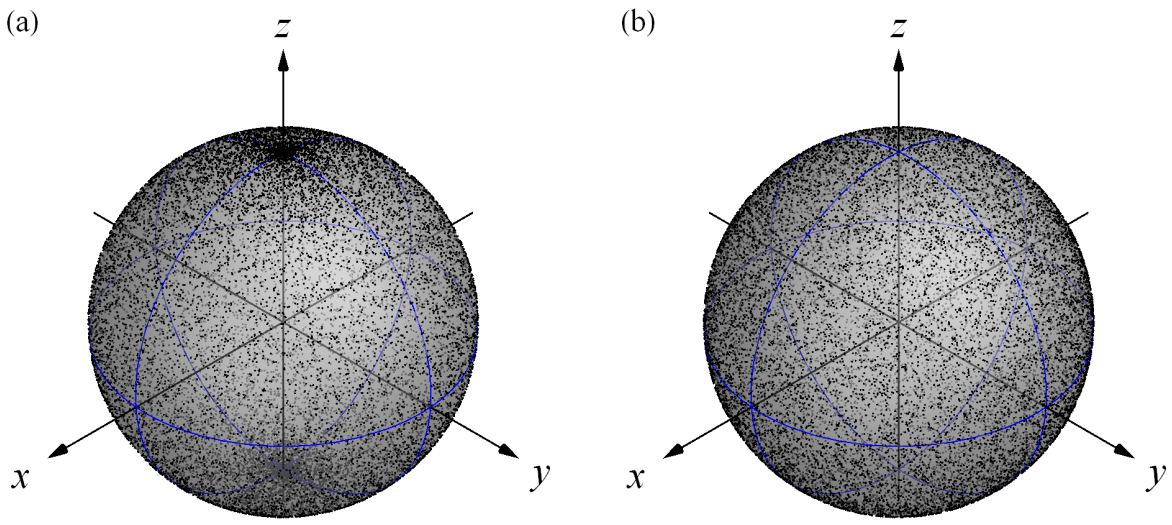

Figure 5.2: Nonuniform (a), and uniform random directions. Each point on the spheres in these figures represents a particular direction.

possible. Depending on the specific particle shape, size, and refractive index, suitable convergence can occur at nearly any number of orientations, assuming that it converges at all. The convergence can determined, for example, from the variation in the averaged extinction cross section over a set of many multi-orientation DDA calculations. Alternatively, one could simply guess at a suitable number of orientations, evaluate Eq. (5.31) for a scattering quantity, then compare to the corresponding quantity obtained from the T-Matrix solution for the same particle. Although less rigorous, this method is vastly faster and will be the method used in Ch. 8.

5.5 Limitations and Verification

The discretization of the VIE (2.10) assumes that the internal field within a volume element ΔV can be represented by a single dipole moment at the element's center. This substitution is strictly possible only if the internal field is exactly uniform throughout the entire element. For time-dependent fields the volume element must then be a point, which requires an infinite number of elements to represent the particle. The necessity of using finite sized elements

in the numerical implementation of the DDA means an inherent error in replacing each element's internal field by a dipole. This discretization error is essentially unavoidable, but vanishes with the element's size. In principle then, the error can always be made negligible with finer discretization.

Shape errors in the DDA concern the inability of the (cubical) discretization lattice to adequately mimic the geometry of the particle's surface. The substantial increase in speed afforded by the FFT-based matrix multiplication involved in the numerical solution Eq. (5.15) essentially makes choosing a cubical discretization lattice mandatory. Although, in principle, one could chose a three-dimensional lattice of any geometry. Nearly any particle shape will have a surface that does not exactly coincide with the geometry of the lattice. The exception, of course, being a cubical particle oriented with its sides parallel to the lattice axes. As with discretization errors, these shape errors vanish with the fineness of the lattice.

Figure (5.3) provides an example of the verification of the DDA programs developed for this work. Here, the DDA is used to calculate the far-field scattered intensity corresponding to a prolate spheroidal particle. The incident wave is linearly polarized along the x-axis, propagates along the z-axis, and the scattered intensity is calculated for points in the y-z plane[7] along the $\mathcal{C}_h$ contour shown in Ch. (6) in Fig. (6.2). The particle is oriented at a 45° angle with respect to the propagation direction of the incident wave, see inset in plot (b). A spheroid is chosen because it can break the symmetry that would otherwise exist in the scattering arrangement involving a spherical particle, see Ch. 6. This provides a more demanding check of the DDA. The size, refractive index, and aspect ratio of the particle are $kR_{ve} = 6.0$, $m = 1.45 + 0.05i$, and $a/b = 0.5$, respectively, and $N = 98732$ discrete volume elements are used in the calculation. The DDA generated scattering curve is compared in plot (a) to the curve obtained from the T-Matrix program of [2, Part II]. Plot (b) shows the root mean square (RMS) deviation between the two curves in (a). Overall, one can see

[7] i.e. the "horizontal scattering plane"

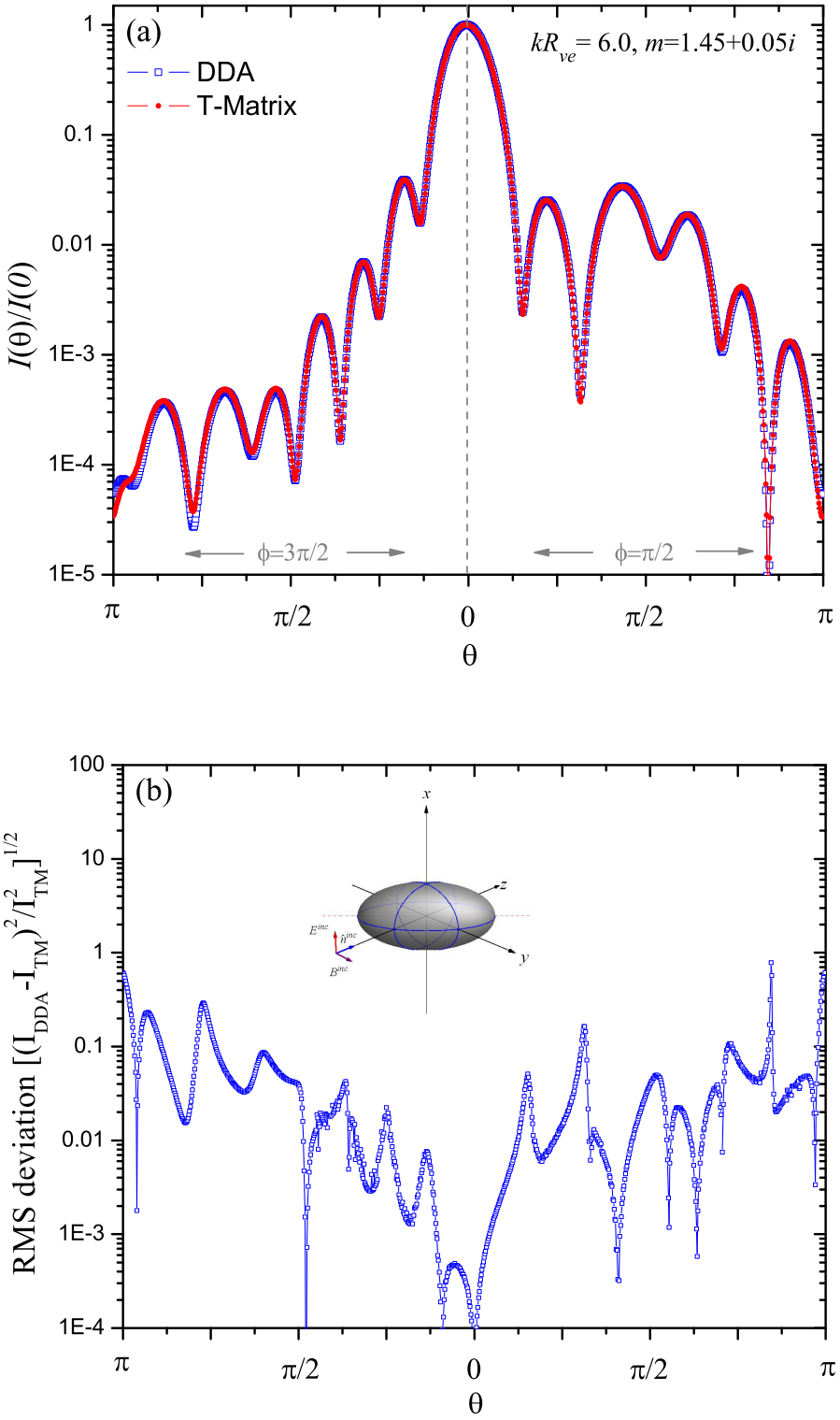

Figure 5.3: T-Matrix verification of the DDA for a prolate spheroidal particle.

good agreement between the DDA and T-Matrix, with errors roughly around $\sim 1-3\%$.

5.6 A Physical Picture

The above discussions primarily focus on the DDA in the context of it being a numerical solution to the VIE (2.10). One could then think that the DDA is important only as a numerical method, and as such, does not provide any insight as to the physical nature of the solution. Indeed, this is often the case with the alternative series-based solutions to the Maxwell equations. However, because the DDA is derived from the VIE, it shares exactly the same physical interpretations that are associated with the microphysical model, recall Ch. 4. Moreover, the intricate VIE-discretization process resulting in the DDA's linear system of equations provides an enhanced understanding of the microphysical model, specifically the concept of wavelet-source coupling.

An intuitive way to describe this "enhanced understanding of the microphysical model," is to visualize the interaction between the dipole's in a *contrived* temporal context. First consider the dipole lattice when the incident plane wave is not present. In this case, the moments have zero magnitude because there is no electric field to polarize them, recall Eq. (5.5). Now suppose that the incident wave is "turned on." Each dipole becomes polarized and oscillates due to the harmonic time-dependence of the incident wave. This oscillation means that a given dipole radiates a secondary wave that propagates through the vacuum filling the lattice and influences the polarization of the other dipoles. The secondary radiation from these other dipoles then act back on the first dipole, and so on, *ad infinitum*. This interaction between the dipoles, enabled by their secondary waves, is called *coupling*. Eventually, a "steady state" is reached, where the dipoles' magnitude, direction, and phase stops evolving under this coupling action and the dipoles collectively attain a state of common time-harmonic oscillation. The end result is that each dipole behaves as if it is driven by the local internal electric field that would otherwise be found by solving the macroscopic Maxwell equations directly, e.g. by the Mie solution of [37, Ch. 4]. Hence,

from this point of view, it is the coupling action that causes the effects that one associates with refraction.

Notice that this microphysical picture is a profoundly different description of electromagnetic scattering than what one typically sees. The concept of a refractive index is essentially abandoned here, and replaced by the requirement that the coupling action between the dipoles results in a self-consistent solution; this is the physical meaning behind the solution to Eq. (5.15). In the alternative approach to solving the Maxwell equations, one simply scales the wavenumber in the particle by the refractive index, i.e., $k \to mk$, in the expansions for the internal field, Eqs (2.20) and (2.23), e.g., see [58, p. 383, footnote 7]. At that point, it is only a matter of matching the boundary conditions at the particle surface to find the expansion coefficients and have a complete solution. It is remarkable then, that the infinitely complicated coupling between the dipoles results in the same effects of reflection, refraction, and scattering, as are also described by the series-based solutions to the Maxwell equations.

The strength of the coupling between the dipoles depends, in part on the polarizability α, and on the angular structure of the dipole fields. The polarizability depends on the refractive index m from Eq. (5.12), and overall, the dipoles are more strongly coupled as the real part of m increases. This brings one to an important realization; *coupling is the cause of refraction* in the microphysical model. If the real part of the refractive index is close to unity, the polarizability is small and the coupling between the dipoles is weak. This means that the incident wave will dominate the polarization of the volume elements, resulting in a situation where the internal field deviates little from the incident wave. This is weak refraction and corresponds to the so-called Rayleigh-Debye-Gans (RDG) approximation of Chs. 7 and 8. If the real part of the refractive index is increased, the coupling also increases and the dipole moments can become substantially deviated from what one would expect from the incident wave only. The result is a refracted internal field. Moreover, when the coupling is weak, the shape of the particle does not influence the polarization of the volume

elements since dipoles near the particle surface do not significantly influence dipoles deeper in the particle[8]. If the coupling is strong, the interaction between these surface and bulk dipoles is enhanced, which can result in a strong shape-dependent influence on the form of the internal field. This is just as one would expect from a simple ray-tracing Snell's law picture of refraction.

Because the coupling is achieved through the dipole's fields, the angular dependence of the near and far zone dipole waves influence the strength of the dipole-dipole coupling, recall Eq. (2.35). If the separation between the dipoles is large compared to the wavelength, the angular dependence has a $\sin^2\theta$ functionality, where θ is measured with respect to the dipole direction. This means that two dipoles with the same direction, which is parallel to the line connecting them, are essentially uncoupled in the each other's far-field zones. In contrast, if these same dipoles are rotated so that they point at right angles to the line connecting them, then they become coupled and remain so even in the far-field.

Elementary texts on classical optics often describe refraction as the "slowing down" of light as it enters a medium from vacuum [34, p. 26], [59, Sec. 4.2]. The speed of the light is reduced to $c/\mathrm{Re}\{\mathrm{m}\}$ and the wavelength is reduced by $\lambda/\mathrm{Re}\{\mathrm{m}\}$. In the context of the microphysical model, where refraction is *caused* by the coupling between a particle's wavelets (or dipoles), one is reminded that *light actually travels inside the particle medium at the same speed as it does through vacuum; the speed of light.* The coupling between the particle's dipoles is achieved by the waves that the dipoles radiate. These waves travel at the vacuum speed c, which can be seen from the appearance of k in the Green function of the VIE (2.10) and the implicit harmonic time dependence $\exp(-i\omega t)$. Recall that the Green function acts as the *free space* propagator of the field. If the dipoles' radiated waves were to travel at the refracted speed $c/\mathrm{Re}\{\mathrm{m}\}$, then $\mathrm{Re}\{\mathrm{m}\}k$ would have to appear in place of k in the Green function.

The *apparent* reduction of the wavelength in a medium can be observed, for example,

[8]This does not mean, however, that the particle shape does not influence the form of the scattered wave, e.g., see the RDG approximation in Ch. 7.

from the effect of thin film interference [59, Sec. 9.4]. So, if the waves exchanged between the wavelets in the medium travel at c, how then is one to account for the observed reduced internal wavelength? A complete answer to this question requires the introduction of the Ewald-Oseen extinction theorem and will be more clear in Sec. 11.5. Nevertheless, an incomplete explanation can be given here. The key is to realize that it is the dipole moments that are the *source* of the scattered wave, which is the wave that is observed. The distribution of moments throughout the particle, i.e., its polarization, has a wavelength shorter than λ, which is due to the shift in phase of the dipole oscillations with respect to the incident wave, see [59, Sec. 4.2.3]. Yet, the coupling interactions that *establish* this polarization are realized by the exchange of dipole-radiated waves of wavelength λ and speed c.

Notice that the description of coupling presented above involves a time-like character; the dipoles are first polarized by the incident wave, they then radiate, influence each other's polarization, and eventually settle into a steady state. One should realize that there really is no separation in actual time between these coupling steps. The reason is because everything in the derivation of the DDA is done in the frequency-domain where the (monochromatic) harmonic time-dependence of all sources and fields is implicit. Because of this, the incident wave is never "on" or "off" and the coupling between the dipoles does not occur in any real temporal sequence. This means that the above discussions relating to refraction do not strictly apply to the propagation of light pulses in a medium. One would have to reformulate the microphysical model to accommodate an incident wave with a finite spectral bandwidth in order to properly describe a transient state where the dipoles become initially polarized as they encounter a pulse. Even so, the coupling interaction between the dipoles will still occur at the speed of light in vacuum.

Chapter 6

Reflection Symmetry and its Consequences

"The fusion of the ideas in geometry and symmetry with physics has proven to be extraordinarily fruitful..."

-C. E. Baum [60]

This chapter applies the microphysical model to calculate and interpret the wave scattered by spherical and nonspherical particles. As a result, a physical picture emerges connecting a particle's reflection symmetry to the polarization state of its scattered wave. This same physical understanding, however, is not evident from the more common approach of analyzing the mathematical character of the series-based solution for the scattered wave, e.g. see [47, Ch. 5], [2, 60–70].

6.1 Spherical Particles

First consider a spherical particle illuminated by the linearly polarized plane wave of Eqs. (3.1) and (3.2), see Fig. (6.1). The particle is surrounded by the same imaginary spherical surface $\mathcal{S}_{en}$ of radius R_{en} as in Ch. 3, except here the size of the surface is large enough that it is in the particle's far-field zone. The observation point $\mathbf{r}$ will be restricted to points on $\mathcal{S}_{en}$ throughout this chapter. Occasionally, reference will be made to the particle's *scattering arrangement*. What is meant by this term is the geometric configuration formed by the

particle and the polarization and propagation direction of the incident wave.

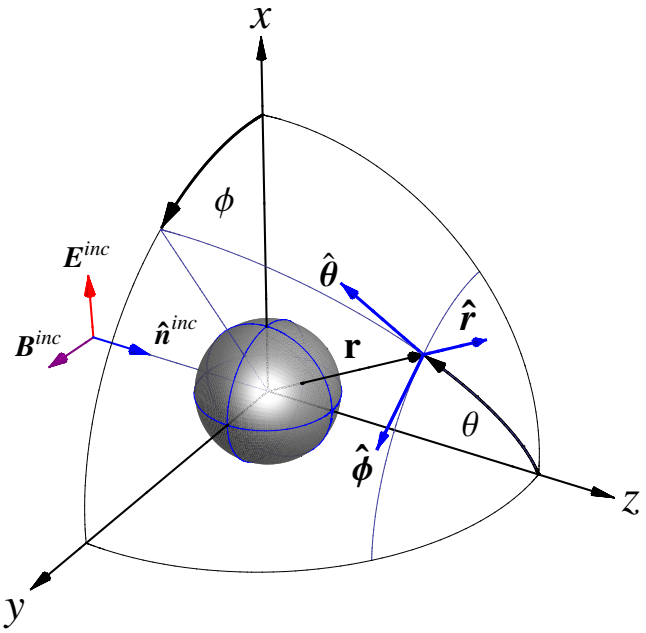

Figure 6.1: Scattering arrangement for a spherical particle.

6.1.1 Symmetry of the Mie Internal Field

The Mie internal electric and magnetic fields presented in [37, Ch. 4] can be cast into a form that readily demonstrates their reflection symmetry. Recall from Eq. (2.21) that any electric field can be expressed in terms of vector spherical wave functions. Doing so for the field inside of a sphere, one has

$$\mathbf{E}^{int}(\mathbf{r}) = \sum_{n=1}^{\infty} E_n \big[c_n \mathbf{M}_n(\mathbf{r}) - i d_n \mathbf{N}_n(\mathbf{r}) \big] \qquad \mathbf{r} \in V^{int}, \tag{6.1}$$

and the magnetic field follows from the Eq. (1.22) as

$$\mathbf{B}^{int}(\mathbf{r}) = -\frac{i}{\omega} \boldsymbol{\nabla} \times \mathbf{E}^{int}(\mathbf{r}) \qquad \mathbf{r} \in V^{int}. \tag{6.2}$$

The connection between $\mathbf{E}^{int}$ and $\mathbf{B}^{int}$ in Eq. (6.2) allows one to consider only the electric field when analyzing the wave's reflection symmetry, see [67]. In Eq. (6.1), $E_n = i^n E_o^{inc}(2n+1)/n(n+1)$ and the functions $\mathbf{M}_n$ and $\mathbf{N}_n$ are not the same as the VSWFs in Sec. 2.2, but are

related to them. These are simplified versions of the more general VSWFs, tailored for the application to spherical particles, and are usually formulated in spherical polar coordinates. Explicit expressions for these functions are given in [37, p. 95]. As usual, the expansion coefficients c_n and d_n in Eq. (6.1) are determined by the boundary conditions at $\mathcal{S}$ and depend only on the sphere's size parameter kR, and refractive index m. Because of their independence on the location of the observation point, c_n and d_n are not explicitly given here, see [37, Eq. (4.52)].

The internal field can be expressed in Cartesian coordinates by defining the radial, polar, and azimuthal functions

$$R_n(\mathbf{r}) = E_n\big[c_n\mathbf{M}_n(\mathbf{r}) - id_n\mathbf{N}_n(\mathbf{r})\big]\cdot\hat{\mathbf{r}}, \tag{6.3}$$

$$\Theta_n(\mathbf{r}) = E_n\big[c_n\mathbf{M}_n(\mathbf{r}) - id_n\mathbf{N}_n(\mathbf{r})\big]\cdot\hat{\boldsymbol{\theta}}, \tag{6.4}$$

$$\Phi_n(\mathbf{r}) = E_n\big[c_n\mathbf{M}_n(\mathbf{r}) - id_n\mathbf{N}_n(\mathbf{r})\big]\cdot\hat{\boldsymbol{\phi}}, \tag{6.5}$$

respectively. Then, the rectangular functions X_n, Y_n and Z_n are defined as

$$X_n(\mathbf{r}) = R_n(\mathbf{r})\,\sin\theta\,\cos\phi + \Theta_n(\mathbf{r})\,\cos\theta\,\cos\phi - \Phi_n(\mathbf{r})\,\sin\phi \tag{6.6}$$

$$Y_n(\mathbf{r}) = R_n(\mathbf{r})\,\sin\theta\,\sin\phi + \Theta_n(\mathbf{r})\,\cos\theta\,\sin\phi + \Phi_n(\mathbf{r})\,\cos\phi, \tag{6.7}$$

$$Z_n(\mathbf{r}) = R_n(\mathbf{r})\,\cos\theta - \Theta_n(\mathbf{r})\,\sin\theta. \tag{6.8}$$

Equations (6.6)-(6.8) now give the internal field in Cartesian coordinates,

$$\mathbf{E}^{int}(\mathbf{r}) = \sum_{n=1}^{\infty}\left[X_n(\mathbf{r})\,\hat{\mathbf{x}} + Y_n(\mathbf{r})\,\hat{\mathbf{y}} + Z_n(\mathbf{r})\,\hat{\mathbf{z}}\right]. \tag{6.9}$$

To describe the reflection symmetry of $\mathbf{E}^{int}$ and the scattering arrangement, let Π_x, Π_y, and Π_z denote the y-z, x-z, and x-y planes through the origin, respectively. The intersection of the large spherical surface $\mathcal{S}_{en}$ with the Π_x and Π_y planes defines the horizontal contour $\mathcal{C}_h$ and the vertical contour $\mathcal{C}_v$, respectively. Note that these contours can have practical significance, since in some simple scattering measurements the detector is confined to one

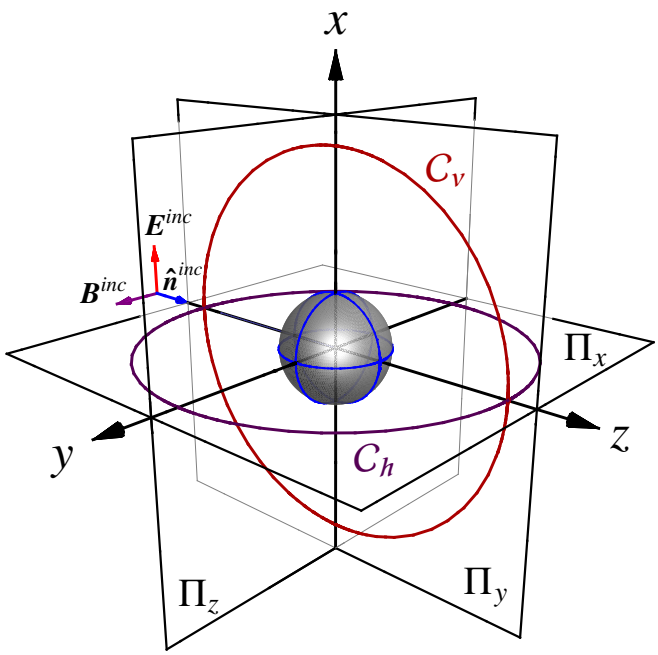

Figure 6.2: The Π_x, Π_y and Π_z planes and $\mathcal{C}_h$ and $\mathcal{C}_v$ contours. The planes pass through the origin and the center of the sphere.

of these contours [71]. Chapter 7 in particular will focus on observation points that reside on $\mathcal{C}_h$. Figure (6.2) shows the configuration of planes and contours, and inspection of this figure demonstrates that the scattering arrangement is invariant under reflection about the Π_x and Π_y planes. It is not however, invariant under reflection about the Π_z plane because of the propagation direction of the incident wave.

The reflection symmetry of $\mathbf{E}^{int}$ is revealed by the behavior of the rectangular functions X_n, Y_n and Z_n under the transformations $x \to -x$ and $y \to -y$, where x, y, and z represent the Cartesian components of $\mathbf{r}$. Upon making these transformations in Eqs. (6.6)-(6.8), one can demonstrate that

$$X_n(x,y,z) = X_n(-x,y,z) = X_n(x,-y,z), \tag{6.10}$$

$$Y_n(x,y,z) = -Y_n(-x,y,z) = -Y_n(x,-y,z), \tag{6.11}$$

$$Z_n(x,y,z) = -Z_n(-x,y,z) = Z_n(x,-y,z). \tag{6.12}$$

Because the components of the internal field must be continuous throughout V^{int}, since no

sources reside in the particle, Eqs. (6.11) and (6.12) provide the additional conditions that

$$Y_n(\mathbf{r}) = 0 \quad \text{for } x = 0 \text{ or } y = 0, \tag{6.13}$$

$$Z_n(\mathbf{r}) = 0 \quad \text{for } x = 0. \tag{6.14}$$

Equation (6.10) shows that the x-component of the internal field is invariant under reflection about the Π_x and Π_y planes, whereas Eq. (6.11) shows that the y-component switches sign upon reflection about these planes. Equation (6.12) demonstrates that the z-component of the field is invariant about the Π_y plane but switches sign about the Π_x plane.

From Eqs. (6.13) and (6.14) one can see that the y-component of the internal field vanishes on the Π_x and Π_y planes and the z-component vanishes for points on the Π_x plane. This means that the internal field can have only an x-component in the Π_x plane, and only the x and z-components on the Π_y plane. Because of Eq. (6.10), the field must have reflection symmetry about the z-axis in the Π_x plane.

To illustrate these symmetries, a numerical implementation of the Mie solution is used to render the internal field for a sphere. Figure (6.3) shows the normalized field magnitude in the Π_x plane for a sphere with $kR = 12$ and $m = 1.33 + 0i$. Figures (6.4) and (6.5) show the same except in the Π_y and Π_z planes. The color code for these plots is given in Fig. (6.3) and is in log scale. The colors are assigned relative to Fig. (6.3), which enables comparison between the plots.

6.1.2 Application of the Microphysical Model

Now apply the microphysical model of Eq. (4.11) to the sphere,

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp(ikr)}{r}\frac{k^2}{4\pi}(m^2-1)\lim_{\Delta V\to 0}\sum_i \left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}_i)\,\exp(-ik\mathbf{r}_i\cdot\hat{\mathbf{r}})\,\Delta V. \tag{6.15}$$

Implicit in Eq. (6.15) is the limit $\Delta V \to 0$, and the sum runs over the locations $\mathbf{r}_i$ of every volume element in the sphere. Next, recall the wavelet phasors of Eqs. (4.28) and (4.29);

$$z_{i\theta}(\hat{\mathbf{r}}) = c_o\left\{\left[\cos\phi\, E_x^{int}(\mathbf{r}_i) + \sin\phi\, E_y^{int}(\mathbf{r}_i)\right]\cos\theta - \sin\theta\, E_z^{int}(\mathbf{r}_i)\right\}\exp(-ik\mathbf{r}_i\cdot\hat{\mathbf{r}}),$$

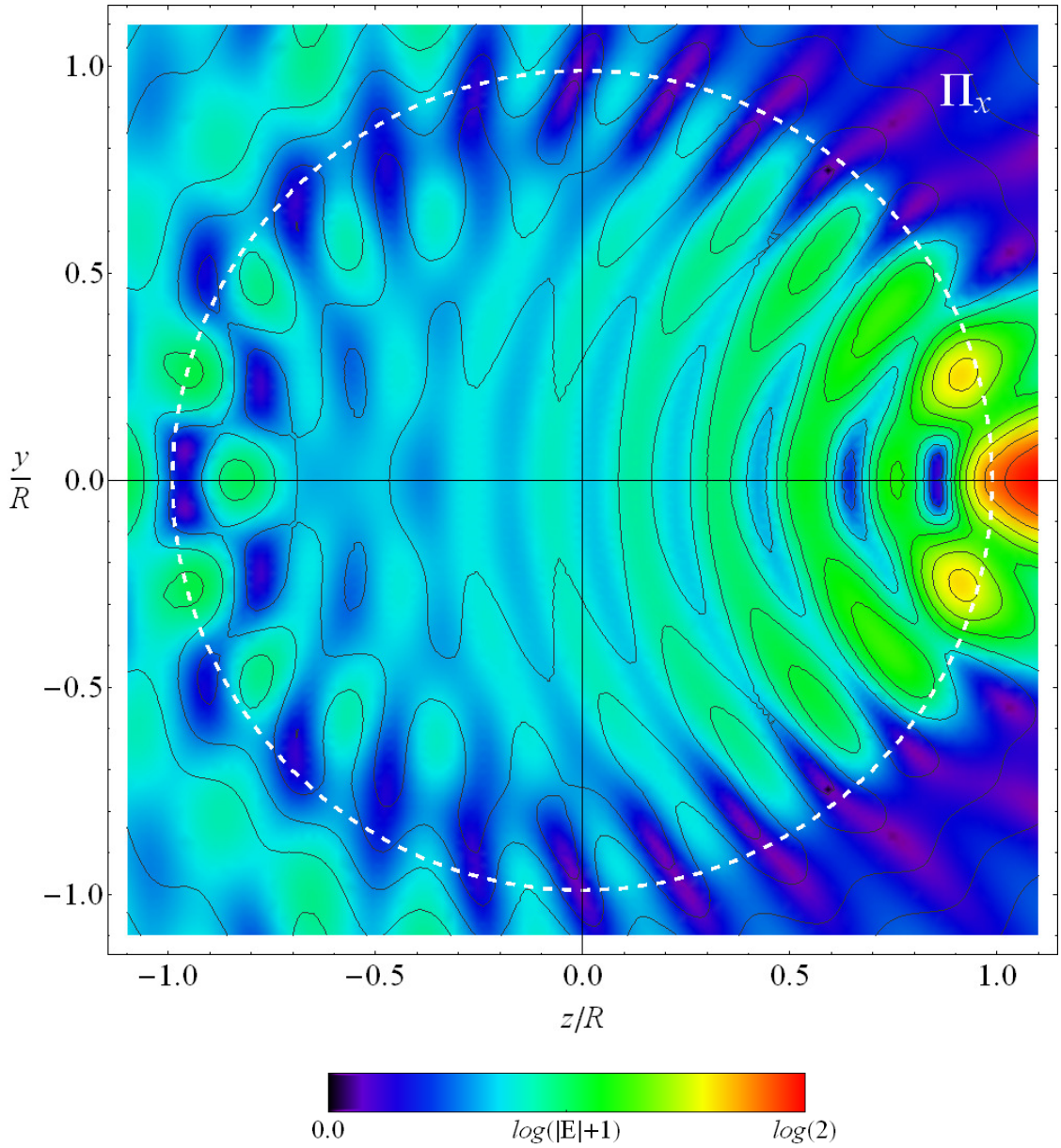

Figure 6.3: Electric field inside and surrounding a sphere with $kR = 12$ and $m = 1.33 + 0i$ for points in the Π_x plane, recall Fig. (6.2). The intersection of the sphere's surface with the Π_x plane is outlined by the white dashed line. The normalized field magnitude is shown by the color shades in log scale. No field vectors are shown since the field is normal to this plane, see Eqs. (6.13) and (6.14).

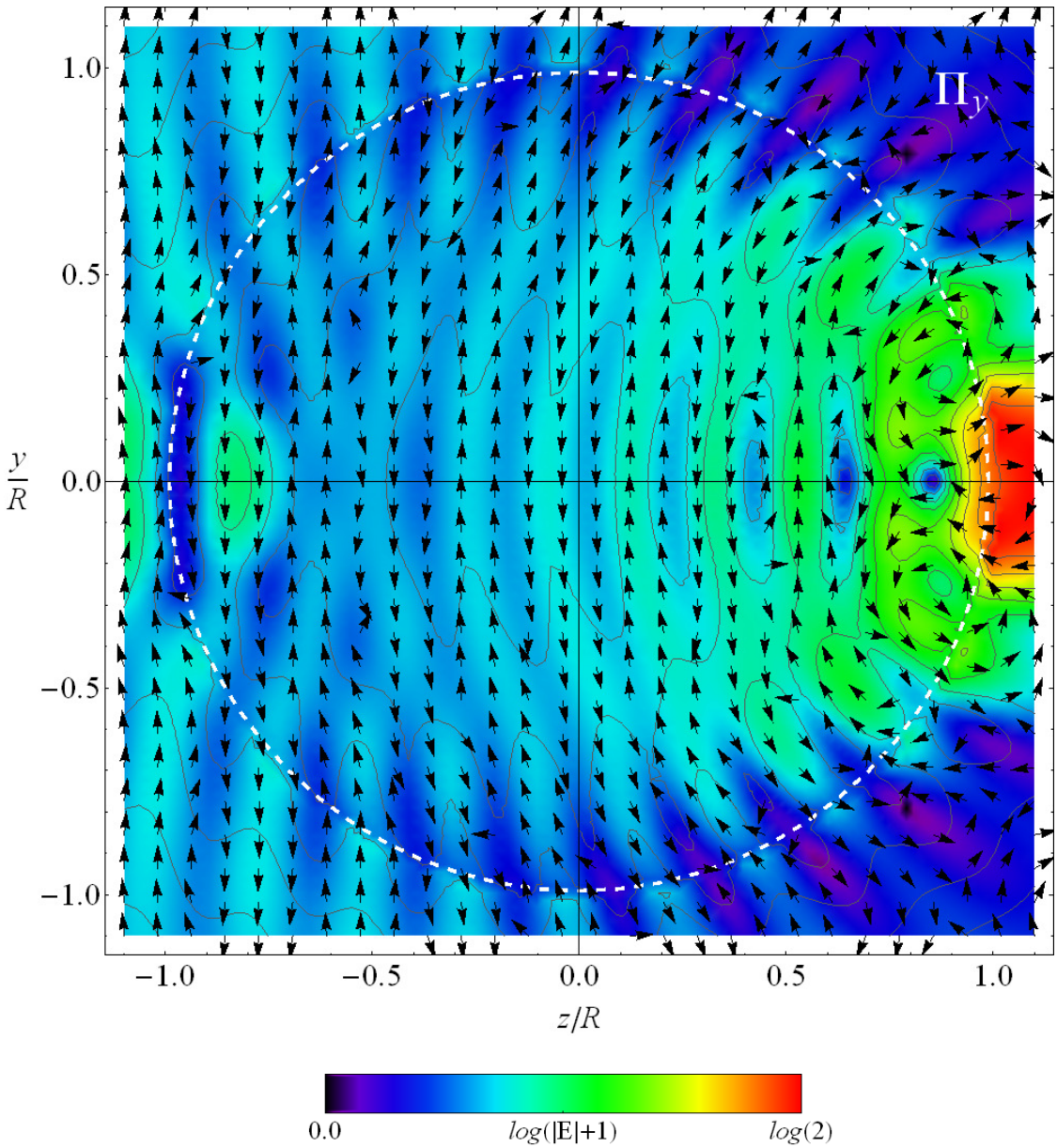

Figure 6.4: Same as Fig. (6.3) except for the Π_y plane.

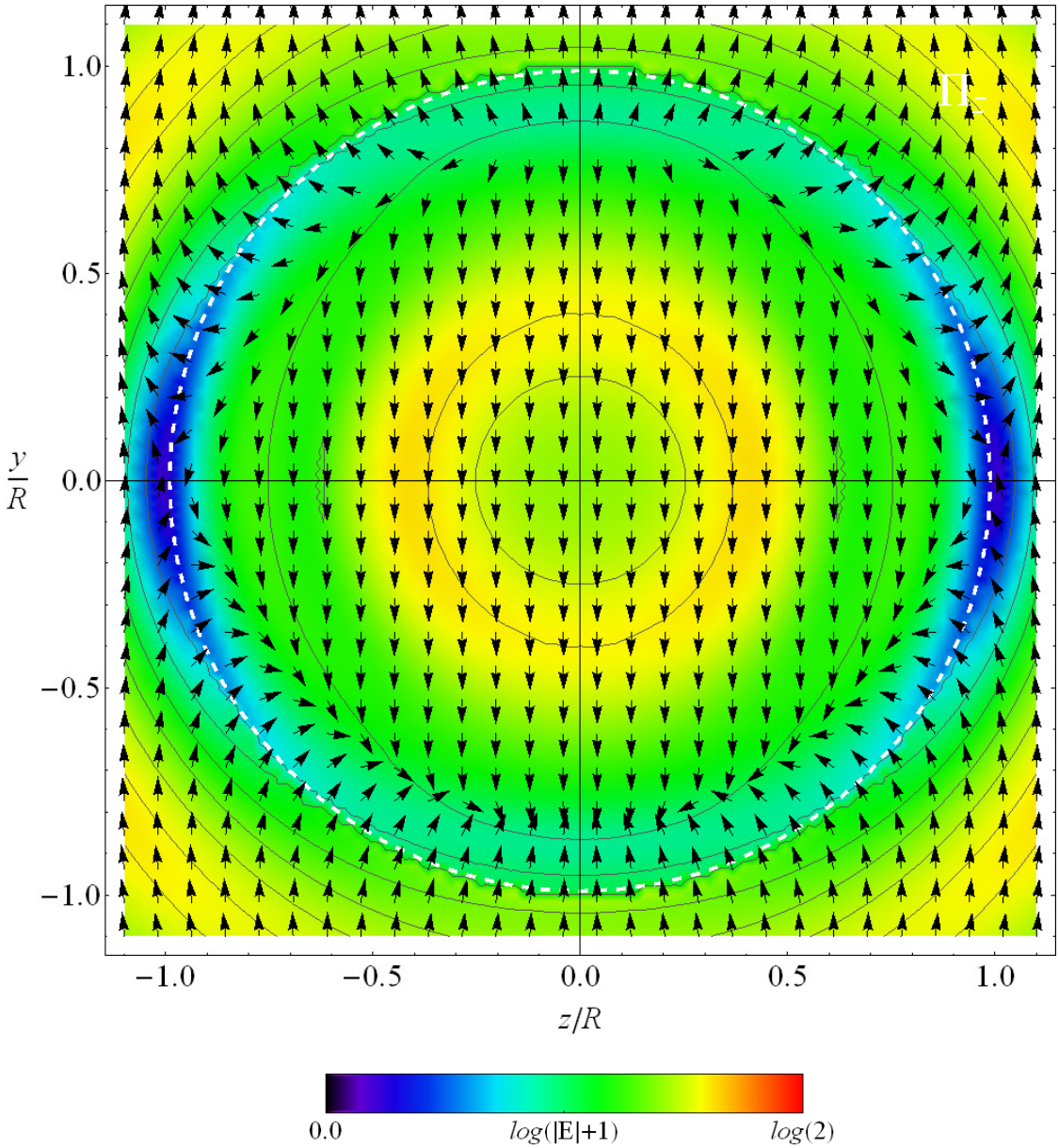

Figure 6.5: Same as Figs. (6.3) and (6.4) except for the Π_z plane.

$$z_{i\phi}(\hat{\mathbf{r}}) = c_o \left[\cos\phi \, E_y^{int}(\mathbf{r}_i) - \sin\phi \, E_x^{int}(\mathbf{r}_i) \right] \exp(-ik\mathbf{r}_i \cdot \hat{\mathbf{r}}),$$

As discussed in Sec. 4.3, the phasors $z_{i\theta}$ and $z_{i\phi}$ represent the amplitudes of the spherical-polar components of the electric field of the i^{th} wavelet. This interpretation is mathematically expressed by Eq. (4.30),

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp(ikr)}{r} \sum_i \left[z_{i\theta}(\hat{\mathbf{r}}) \, \hat{\boldsymbol{\theta}} + z_{i\phi}(\hat{\mathbf{r}}) \, \hat{\boldsymbol{\phi}} \right].$$

Now place the observation point somewhere on the $\mathcal{C}_h$ contour. The azimuthal angle can have only two values, $\phi = \pi/2$ or $\phi = 3\pi/2$. One can then reduce Eq. (4.28) to

$$z_{i\theta}(\hat{\mathbf{r}}) = -c_o \left[\sin\theta \, E_z^{int}(\mathbf{r}_i) \mp \cos\theta \, E_y^{int}(\mathbf{r}_i) \right] \exp(-ik\mathbf{r}_i \cdot \hat{\mathbf{r}}) \,, \tag{6.16}$$

and Eq. (4.29) to

$$z_{i\phi}(\hat{\mathbf{r}}) = \mp c_o \, E_x^{int}(\mathbf{r}_i) \exp(-ik\mathbf{r}_i \cdot \hat{\mathbf{r}}) \,, \tag{6.17}$$

where the top and bottom signs in Eq. (6.16) correspond to $\phi = \pi/2$ and $\phi = 3\pi/2$, respectively. Next, the sum over the sphere's volume in Eq. (4.30) for the θ-component of the field is split in two. One sum covers the particle hemisphere containing internal points with $x \geqslant 0$ and the other sum covers the hemisphere with $x < 0$, i.e.,

$$\sum_i z_{i\theta}(\hat{\mathbf{r}}) = \sum_j^{x\geqslant} z_{j\theta}(\hat{\mathbf{r}}) + \sum_k^{x<} z_{k\theta}(\hat{\mathbf{r}}) \,, \tag{6.18}$$

where the points $\mathbf{r}_j$ and $\mathbf{r}_k$ are mirror points about the Π_x plane, see Fig. (6.6). From Eqs. (6.11) and (6.12) one sees that the y and z-components of the internal field appearing in Eq. (6.16) change sign upon reflection about the Π_x plane. This means that the sums over the $x \geqslant 0$ and $x < 0$ hemispheres in Eq. (6.18) *cancel each other term-by-term*, hence there is complete destructive interference,

$$z_{j\theta}(\hat{\mathbf{r}}) = -z_{k\theta}(\hat{\mathbf{r}}). \tag{6.19}$$

From Eqs. (4.30) and (6.17), then,

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \mp c_o \frac{\exp(ikR_{en})}{R_{en}} \sum_i E_x^{int}(\mathbf{r}_i) \exp(-ik\mathbf{r}_i \cdot \hat{\mathbf{r}}) \, \hat{\boldsymbol{\phi}} \quad \mathbf{r} \in \mathcal{C}_h \,, \tag{6.20}$$

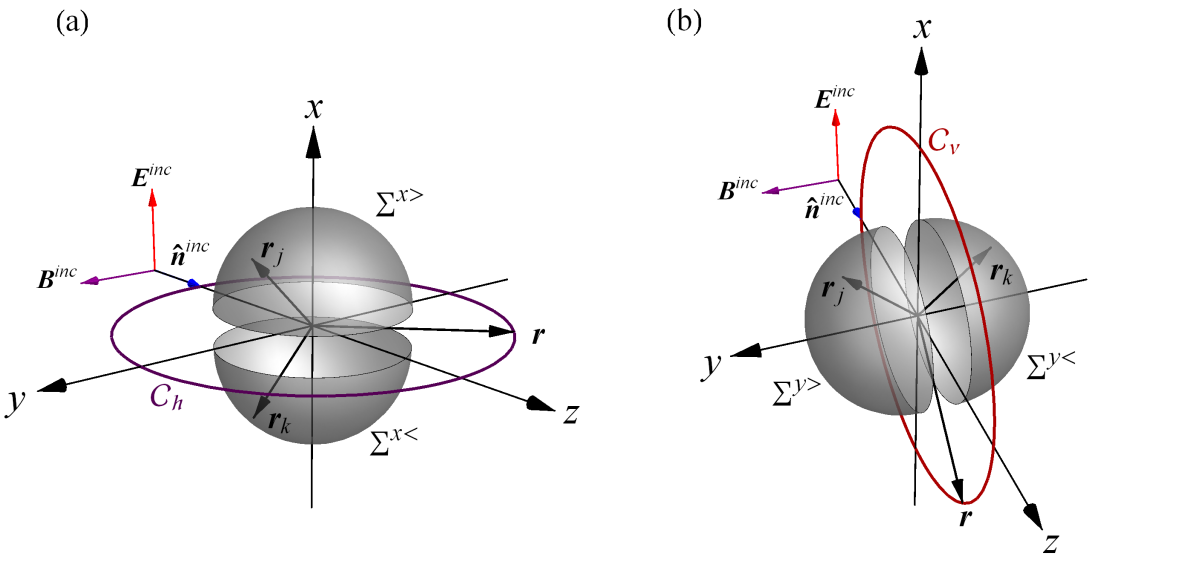

Figure 6.6: Hemispheres used to derive Eqs. (6.20) and (6.25). The sphere's surface is separated to delineate the hemispheres. The internal locations $\mathbf{r}_j$ and $\mathbf{r}_k$ are mirror points about the Π_x plane (a), and the Π_y plane (b).

where the top and bottom signs correspond to $\phi = \pi/2$ and $\phi = 3\pi/2$, respectively.

Equation (6.20) reveals that *for observation points on the $\mathcal{C}_h$ contour the scattered wave is determined solely by the x-component of the sphere's internal electric field.* The symmetry of the internal field leads to the cancellation of the θ-component of the wavelet amplitudes, which carry the dependence on the other Cartesian components of the internal field. Moreover, the direction of the scattered field for observation points on this contour is only along the $\hat{\boldsymbol{\phi}}$ direction[1], thus demonstrating that *the scattered wave for points on $\mathcal{C}_h$ has the same (linear) polarization state as the incident wave.* One will see later that this is not generally true for points on $\mathcal{S}_{en}$.

Next consider placing the observation point on the $\mathcal{C}_v$ contour. Here, $\phi = 0$ and $\phi = \pi$. Equations (4.28) and (4.29) become

$$z_{i\theta}(\hat{\mathbf{r}}) = -c_o \left[\sin\theta \, E_z^{int}(\mathbf{r}_i) \mp \cos\theta \, E_x^{int}(\mathbf{r}_i) \right] \exp(-ik\mathbf{r}_i \cdot \hat{\mathbf{r}}), \tag{6.21}$$

and

$$z_{i\phi}(\hat{\mathbf{r}}) = \pm c_o \, E_y^{int}(\mathbf{r}_i) \exp(-ik\mathbf{r}_i \cdot \hat{\mathbf{r}}), \tag{6.22}$$

[1] which is parallel or anti-parallel to the $\hat{\mathbf{x}}$ direction.

where the top and bottom signs in Eqs. (6.21) and (6.22) correspond to $\phi = 0$ and $\phi = \pi$, respectively.

The sum in Eq. (4.30) is again split in two, as above, except now the hemispheres correspond to $y \geqslant 0$ and $y < 0$, see Fig. (6.6). Then, for the ϕ-component phasors,

$$\sum_i z_{i\phi}(\hat{\mathbf{r}}) = \sum_j^{y\geqslant} z_{j\phi}(\hat{\mathbf{r}}) + \sum_k^{y<} z_{k\phi}(\hat{\mathbf{r}})\,, \tag{6.23}$$

where $\mathbf{r}_j$ and $\mathbf{r}_k$ are now mirror points about the Π_y plane. From the behavior of the y-component of $\mathbf{E}^{int}$ given in Eq. (6.11) combined with Eq. (6.22), one finds that

$$z_{j\phi}(\hat{\mathbf{r}}) = -z_{k\phi}(\hat{\mathbf{r}}), \tag{6.24}$$

which, from Eqs. (4.30) and (6.22), demonstrates that *there is no contribution to the scattered field on $\mathcal{C}_v$ from the y-component of the internal electric field.* Then,

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = -c_o \frac{\exp(ikR_{en})}{R_{en}} \sum_i \left[\sin\theta\, E_z^{int}(\mathbf{r}_i) \mp \cos\theta\, E_x^{int}(\mathbf{r}_i)\right] \exp(-ik\mathbf{r}_i\cdot\hat{\mathbf{r}})\,\hat{\boldsymbol{\theta}}\,, \qquad \mathbf{r}\in\mathcal{C}_v, \tag{6.25}$$

where again the top and bottom signs correspond to $\phi = 0$ and $\phi = \pi$, respectively. This result demonstrates that *the scattered electric field on $\mathcal{C}_v$ has only a θ-component and hence is linearly polarized.*

Using the Stokes parameter formalism of Sec. 1.6, the polarization state of the far-field scattered wave is expressed below in terms of the wavelet amplitudes of Eqs. (4.28) and (4.29). In the process of doing this, one obtains a double sum over the sphere's volume for each Stokes parameter. To explicitly build the symmetries of the sphere's internal field into these parameters, the double sums, which cover the entire sphere volume, are separated into component sums of the volume elements contained in the four regions of the sphere's interior bounded by the intersection of the Π_x and Π_y planes and $\mathcal{S}$. These wedge-shaped regions, or quadrants for short, are labeled by the Roman numerals I-IV, and the volume of each quadrant is denoted by V^{I}-V^{IV}.

Using the symmetry properties of Eqs. (6.10)-(6.12), the double sums over the sphere's four quadrants are reduced to a single double sum over *only the first* quadrant V^{I}. To

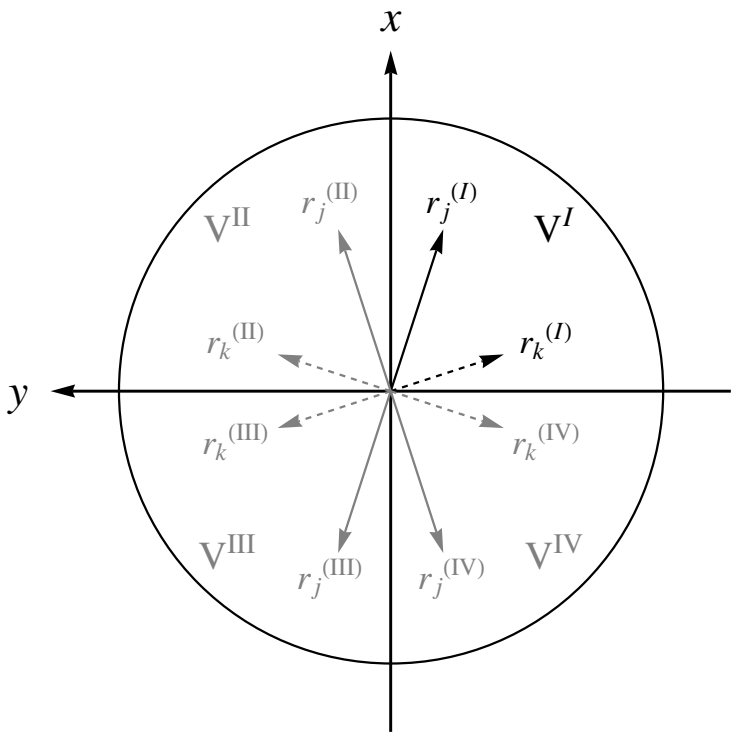

Figure 6.7: Internal points and quadrants of the sphere used to derive Eqs. (6.26)-(6.29).

explain how this is done, consider two arbitrary volume element locations $\mathbf{r}_j^{(\mathrm{I})}$ and $\mathbf{r}_k^{(\mathrm{I})}$ in V^{I}. Let $\mathbf{r}_j^{(\mathrm{II})}$, $\mathbf{r}_j^{(\mathrm{III})}$, $\mathbf{r}_j^{(\mathrm{IV})}$ and $\mathbf{r}_k^{(\mathrm{II})}$, $\mathbf{r}_k^{(\mathrm{III})}$, $\mathbf{r}_k^{(\mathrm{IV})}$ be the locations obtained in the remaining three quadrants by performing successive reflections of $\mathbf{r}_j$ and $\mathbf{r}_k$ about the Π_x and Π_y planes. Figure (6.7) shows these locations in their respective quadrants. Next, let $z_{j\theta}^{(\mathrm{I})}$, $z_{j\phi}^{(\mathrm{I})}$, $z_{k\theta}^{(\mathrm{I})}$, $z_{k\phi}^{(\mathrm{I})}$ be the wavelet amplitudes of Eqs. (4.28) and (4.29) evaluated at the points $\mathbf{r}_j^{(\mathrm{I})}$ and $\mathbf{r}_k^{(\mathrm{I})}$, respectively. Lastly, let $z_{j\theta}^{(\mathrm{II})}$, $z_{j\phi}^{(\mathrm{II})}$, $\ldots$ $z_{j\theta}^{(\mathrm{IV})}$, $z_{j\phi}^{(\mathrm{IV})}$ be the wavelet amplitudes $z_{j\theta}$ and $z_{j\phi}$ evaluated at the reflected points $\mathbf{r}_j^{(\mathrm{II})}$, $\mathbf{r}_j^{(\mathrm{III})}$, $\mathbf{r}_j^{(\mathrm{IV})}$, and similarly for the k subscript terms. Because the points in the V^{II}-V^{IV} quadrants are related to the points in the V^{I} quadrant through the same reflection operations that define the reflection symmetry of the internal field, the wavelet amplitudes in the V^{II}-V^{IV} quadrants are directly related to the amplitudes in only the V^{I} quadrant. With these considerations, the Stokes parameters are expressed Eqs.

(6.26)-(6.29) below as double sums over only the $V^{\rm I}$ quadrant,

$$I(\hat{\mathbf{r}}) = \frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\sum_{j,k}^{(\rm I)}\sum_{\alpha,\beta}\Big\{z_{j\theta}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\theta}^{(\beta)}(\hat{\mathbf{r}})\big]^* + z_{j\phi}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\phi}^{(\beta)}(\hat{\mathbf{r}})\big]^*\Big\}, \tag{6.26}$$

$$Q(\hat{\mathbf{r}}) = \frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\sum_{j,k}^{(\rm I)}\sum_{\alpha,\beta}\Big\{z_{j\theta}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\theta}^{(\beta)}(\hat{\mathbf{r}})\big]^* - z_{j\phi}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\phi}^{(\beta)}(\hat{\mathbf{r}})\big]^*\Big\}, \tag{6.27}$$

$$U(\hat{\mathbf{r}}) = -\frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\sum_{j,k}^{(\rm I)}\sum_{\alpha,\beta}\Big\{z_{j\theta}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\phi}^{(\beta)}(\hat{\mathbf{r}})\big]^* + z_{j\phi}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\theta}^{(\beta)}(\hat{\mathbf{r}})\big]^*\Big\}, \tag{6.28}$$

$$V(\hat{\mathbf{r}}) = \frac{i}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\sum_{j,k}^{(\rm I)}\sum_{\alpha,\beta}\Big\{z_{j\phi}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\theta}^{(\beta)}(\hat{\mathbf{r}})\big]^* - z_{j\theta}^{(\alpha)}(\hat{\mathbf{r}})\big[z_{k\phi}^{(\beta)}(\hat{\mathbf{r}})\big]^*\Big\}, \tag{6.29}$$

where α, $\beta \in$ {I,II,III,IV}, and the notation $\sum^{(\rm I)}$ means that the sum covers the first quadrant $V^{\rm I}$ only. The explicit forms for the wavelet amplitudes $z_{j\theta}^{(\alpha)}$ and $z_{j\phi}^{(\alpha)}$ of Eqs. (6.26)-(6.35) required to prove the symmetry properties of the Stokes parameters in Eqs. (6.30)-(6.33) are:

$$z_{j\theta}^{(\rm I)}(\hat{\mathbf{r}}) = c_o\Big[E_x(\mathbf{r}_j)\,\cos\theta\cos\phi + E_y(\mathbf{r}_j)\,\cos\theta\sin\phi - E_z(\mathbf{r}_j)\,\sin\theta\Big] \times \exp\Big[-ik\big(r_{jx}\sin\theta\cos\phi + r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

$$z_{j\theta}^{(\rm II)}(\hat{\mathbf{r}}) = c_o\Big[E_x(\mathbf{r}_j)\,\cos\theta\cos\phi - E_y(\mathbf{r}_j)\,\cos\theta\sin\phi - E_z(\mathbf{r}_j)\,\sin\theta\Big] \times \exp\Big[-ik\big(r_{jx}\sin\theta\cos\phi - r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

$$z_{j\theta}^{(\rm III)}(\hat{\mathbf{r}}) = c_o\Big[E_x(\mathbf{r}_j)\,\cos\theta\cos\phi + E_y(\mathbf{r}_j)\,\cos\theta\sin\phi + E_z(\mathbf{r}_j)\,\sin\theta\Big] \times \exp\Big[-ik\big(-r_{jx}\sin\theta\cos\phi - r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

$$z_{j\theta}^{(\rm IV)}(\hat{\mathbf{r}}) = c_o\Big[E_x(\mathbf{r}_j)\,\cos\theta\cos\phi - E_y(\mathbf{r}_j)\,\cos\theta\sin\phi + E_z(\mathbf{r}_j)\,\sin\theta\Big] \times \exp\Big[-ik\big(-r_{jx}\sin\theta\cos\phi + r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

and

$$z_{j\phi}^{(\rm I)}(\hat{\mathbf{r}}) = c_o\Big[E_y(\mathbf{r}_j)\,\cos\phi - E_x(\mathbf{r}_j)\,\sin\phi\Big] \times \exp\Big[-ik\big(r_{jx}\sin\theta\cos\phi + r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

$$z_{j\phi}^{(\mathrm{II})}(\hat{\mathbf{r}}) = c_o\Big[-E_y(\mathbf{r}_j)\,\cos\phi - E_x(\mathbf{r}_j)\,\sin\phi\Big] \times \exp\Big[-ik\big(r_{jx}\sin\theta\cos\phi - r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

$$z_{j\phi}^{(\mathrm{III})}(\hat{\mathbf{r}}) = c_o\Big[E_y(\mathbf{r}_j)\,\cos\phi - E_x(\mathbf{r}_j)\,\sin\phi\Big] \times \exp\Big[-ik\big(-r_{jx}\sin\theta\cos\phi - r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

$$z_{j\phi}^{(\mathrm{IV})}(\hat{\mathbf{r}}) = c_o\Big[-E_y(\mathbf{r}_j)\,\cos\phi - E_x(\mathbf{r}_j)\,\sin\phi\Big] \times \exp\Big[-ik\big(-r_{jx}\sin\theta\cos\phi + r_{jy}\sin\theta\sin\phi + r_{jz}\cos\theta\big)\Big],$$

where $c_o = k^2(m^2-1)\Delta V/(4\pi)$, and $\mathbf{r}_j = r_{jx}\,\hat{\mathbf{x}} + r_{jy}\,\hat{\mathbf{y}} + r_{jz}\,\hat{\mathbf{z}}$. These expressions are identical for the k subscript terms.

By making the transformations $\phi \to \pi - \phi$ and $\phi \to 2\pi - \phi$ in Eqs. (6.26)-(6.29), which describe the reflection of the *observation point* about the Π_x and Π_y planes, respectively, one finds that

$$I(\theta,\phi) = I(\theta,\pi-\phi) = I(\theta,2\pi-\phi), \tag{6.30}$$

$$Q(\theta,\phi) = Q(\theta,\pi-\phi) = Q(\theta,2\pi-\phi), \tag{6.31}$$

$$U(\theta,\phi) = -U(\theta,\pi-\phi) = -U(\theta,2\pi-\phi), \tag{6.32}$$

$$V(\theta,\phi) = -V(\theta,\pi-\phi) = -V(\theta,2\pi-\phi). \tag{6.33}$$

The symmetry conditions of Eqs. (6.30)-(6.33) result from the *term-by-term* equality of Eqs. (6.26)-(6.29) under the reflection transformations. That this is so is a direct consequence of the internal field's symmetry about these same planes. Equations (6.32) and (6.33) and the requisite angular continuity of the Stokes parameters provides the extra conditions

$$U(\theta,0) = U(\theta,\pi/2) = U(\theta,\pi) = U(\theta,3\pi/2) = 0, \tag{6.34}$$

$$V(\theta,0) = V(\theta,\pi/2) = V(\theta,\pi) = V(\theta,3\pi/2) = 0. \tag{6.35}$$

It is important to stress that these symmetries of the Stokes parameters require *all* the reflection symmetries of the spheres internal field. Equations (6.30)-(6.35) are also obtained

by [47], [63, 67], however, via different methods; there the analysis concerns either the scattered field's mathematical structure only, or the elements of a matrix transformation of the incident wave into the scattered wave. Neither demonstrate the effect of the internal wave's structure directly.

Notice that if the internal electric field is uniform and directed along the x-axis, then from Eqs. (4.28), (4.29) and (6.29) one finds that

$$V(\hat{\mathbf{r}}) = 0 \quad \text{for} \quad \mathbf{E}^{int}(\mathbf{r}) = E_x^{int}\,\hat{\mathbf{x}}, \tag{6.36}$$

where E_x^{int} is a constant. A situation that would produce such a uniform internal field would be the so-called Rayleigh limit [37]. Equation (6.36) demonstrates that the scattered wave is linearly polarized in all directions, which is consistent with the Rayleigh limit[2], recall Sec. 1.6. Following the wavelet-component cancellation scheme described in Sec. 6.1.4 below, one will be able to understand how this Rayleigh-limit behavior can be generalized to apply to an internal field pointing in *any* direction. This means that *any particle, of any shape, which has a uniform unidirectional internal field, will radiate a scattered wave with the same polarization state as the incident wave in all directions.*

6.1.3 Affect of the Symmetry on the Polarization State

The polarization state of the scattered wave for a sphere with $kR = 4$ and $m = 1.25 + 0i$ is examined in detail below using Mie theory to illustrate the symmetries presented above. The size and refractive index of the sphere is chosen such that the structures in the scattered wave are not too complicated to examine graphically. Other kR and m values have been investigated and give the same qualitative results.

Figure (6.8) shows the polarization state of the scattered wave on $\mathcal{S}_{en}$ where the ellipticity is indicated by gray shades. Lighter shades correspond to more linear polarization whereas darker shades correspond to more circular polarization. The polarization ellipses are also

[2]Referring to [24, p. 780] shows that this result can, of course, also be deduced from the Mie series directly.

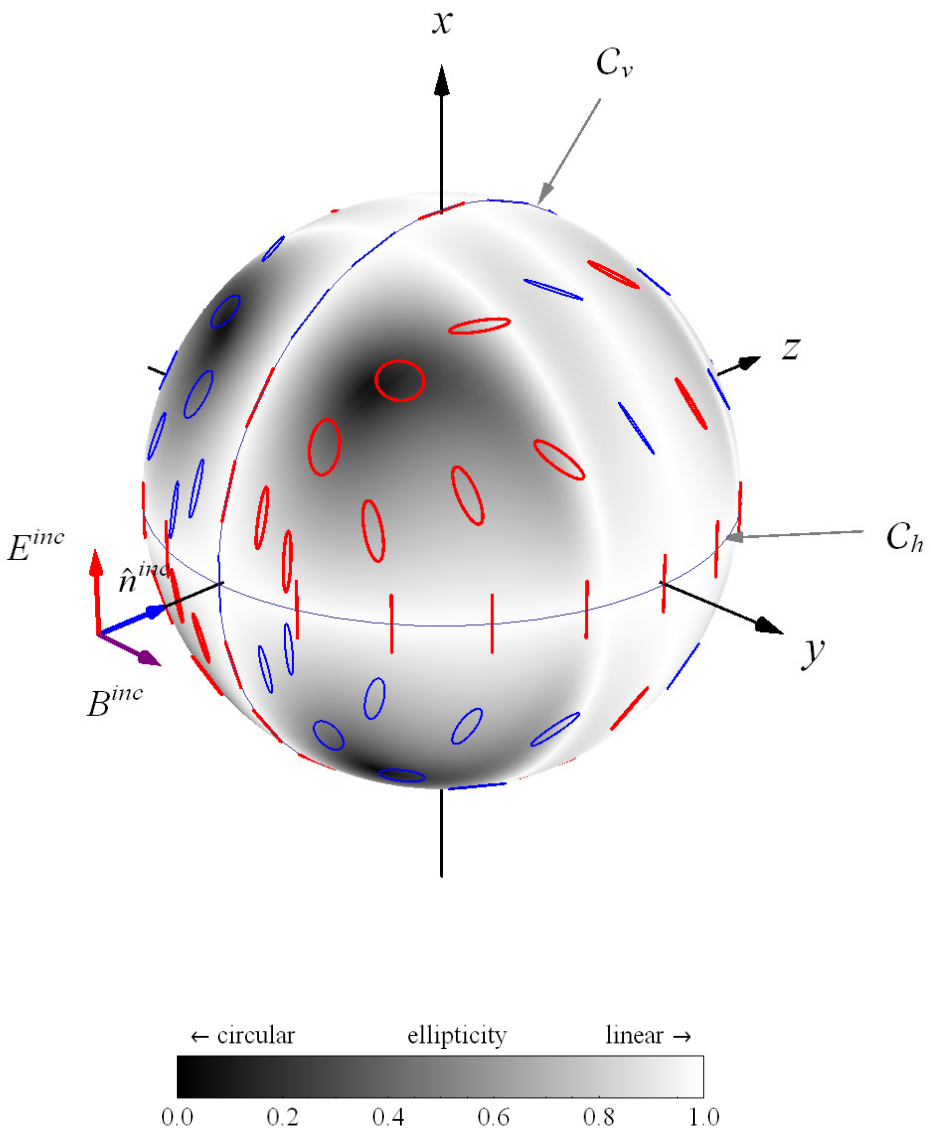

Figure 6.8: Polarization state of the scattered wave for a sphere with $kR = 4$ and $m = 1.25 + 0i$. The plot shows the vibration ellipses at various points on $\mathcal{S}_{en}$, recall Fig. (6.1). The bold red ellipses correspond to right handed rotation of the field whereas thin blue ellipses correspond to left rotation. The ellipticity is shown in gray shades on $\mathcal{S}_{en}$. Darker (lighter) shades indicate more elliptical (linear) polarization.

shown at various points on $\mathcal{S}_{en}$. The right-handed rotation of an ellipse is indicated by a thick red line and the left-handed rotation is indicated by a thin blue line, recall Sec. 1.6. Inspection of this figure shows that the wave is linearly polarized along the $\mathcal{C}_h$ and $\mathcal{C}_v$ contours as predicted by Eqs. (6.20) and (6.25).

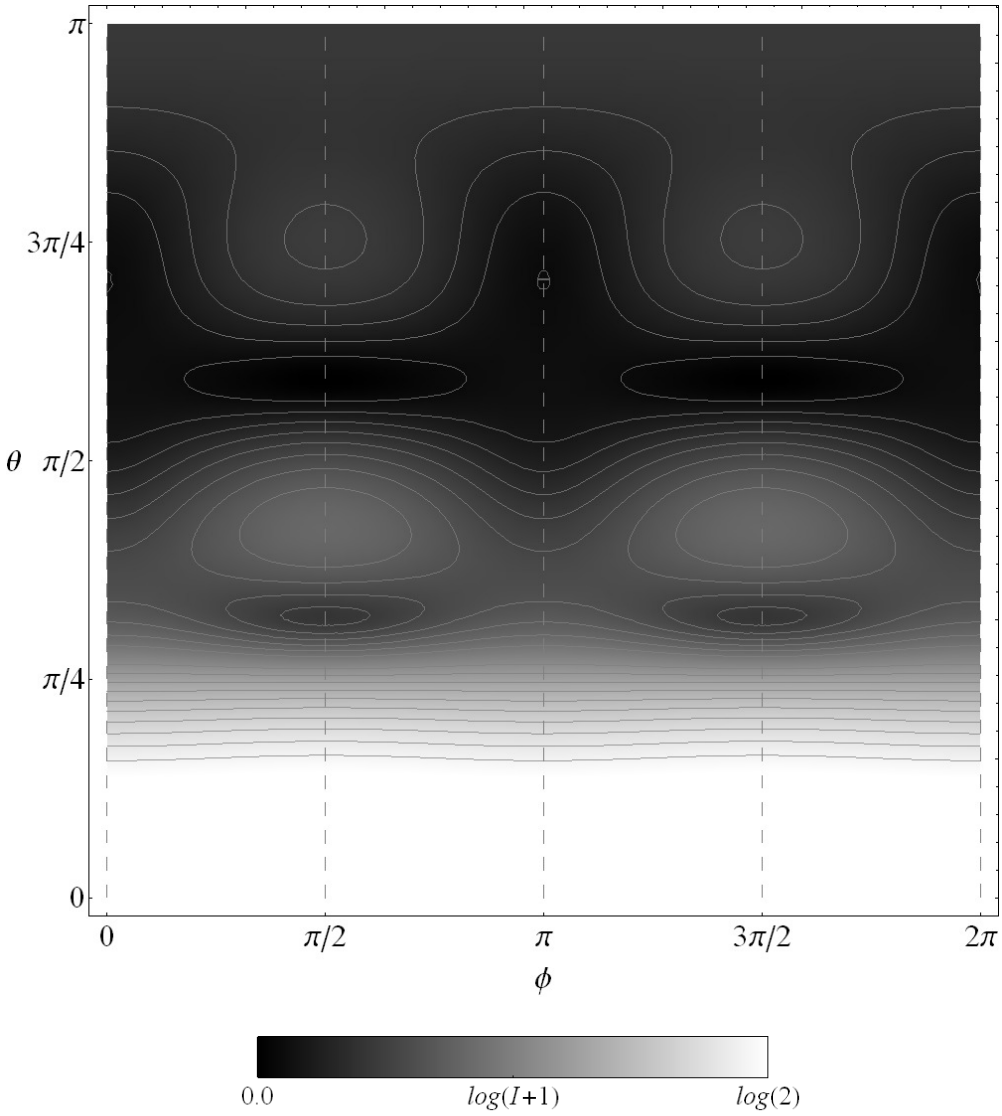

Figure 6.9: Intensity of the scattered wave on $\mathcal{S}_{en}$. The intensity is normalized to the forward direction ($\theta = 0$), and gray shades are assigned in log scale as indicated. The sphere is the same as in Fig. (6.8).

Figures (6.9)-(6.11) display the intensity, ellipticity and rotation of the scattered wave on $\mathcal{S}_{en}$ for the same sphere as in Fig. (6.8), except here the plots show all directions. The

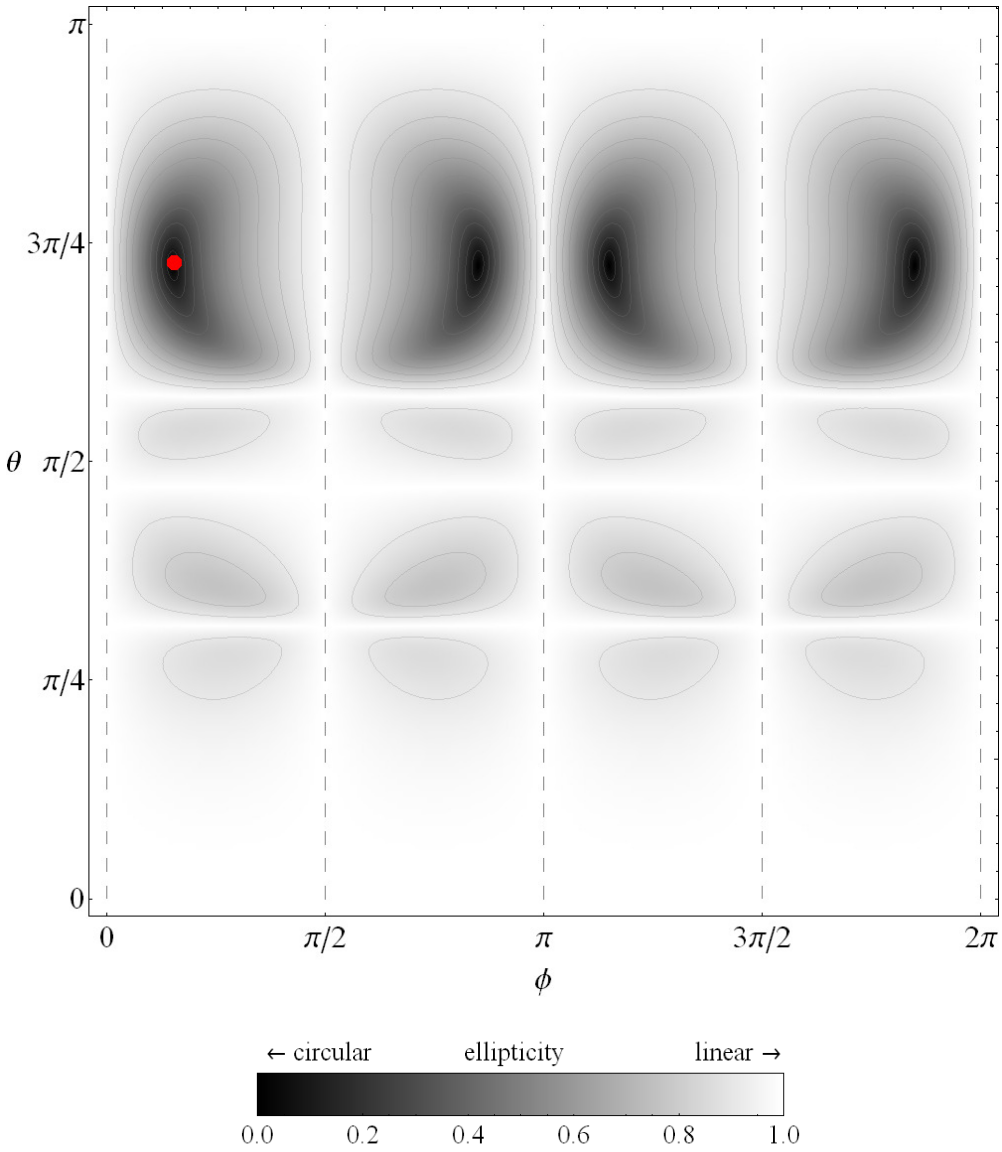

Figure 6.10: Ellipticity of the scattered wave on $\mathcal{S}_{en}$. The gray scale is the same as in Fig. (6.8) and the sphere is the same as in Figs. (6.8) and (6.9). The dashed lines show the intersection of the Π_x and Π_y planes with $\mathcal{S}_{en}$. These lines also indicate the location of four of the seven $\mathcal{L}$-lines, that separate the the angular regions of opposing rotation in Fig. (6.11). One of the four points of circular polarization, or $\mathcal{C}$-points, is indicated by the red dot in the upper left of the plot. See the text for a discussion of $\mathcal{L}$-lines and $\mathcal{C}$-points.

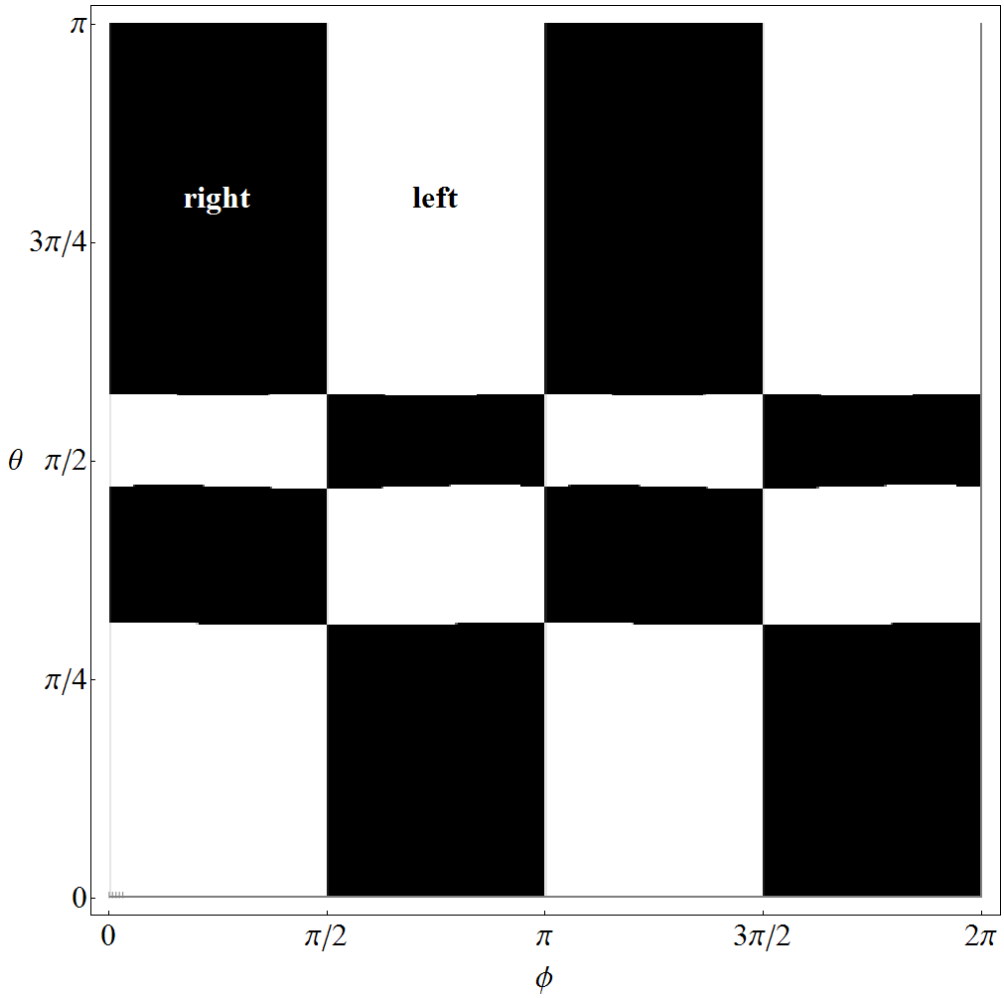

Figure 6.11: Rotation of the scattered wave on $\mathcal{S}_{en}$. Black indicates left-handed rotation and white indicates right-handed rotation. The sphere is the same as in Figs. (6.8)-(6.10). Notice the correlation of the angular structure with that of Fig. (6.10).

intersection of the Π_x and Π_y planes with $\mathcal{S}_{en}$ are indicated on the plots by dashed lines. The intensity, Fig. (6.9), shows reflection symmetry about both the Π_x and Π_y planes. Notice from Fig. (6.10) *that the scattered wave is, in general, elliptically polarized in all directions except those contained in the* Π_x *and* Π_y *planes.* The ellipticity displays reflection symmetry about the Π_x and Π_y planes and shows that *the wave is, overall, more linearly polarized in the forward directions as compared to the backward directions.* The plot of the wave's rotation in Fig. (6.11) reveals that the wave changes its sense of rotation upon reflection of the observation point about either of these planes. The angular distribution of the wave's rotation appears to be "conserved" in the sense that for each angular region of right-handed rotation, there is another region that is the mirror image to that region about the Π_x or Π_y planes that shows left-handed rotation.

Every reflection symmetry of Fig. (6.8)-(6.11) is now explained in the context of the microphysical model. The figures demonstrate that the scattered wave is linearly polarized along the $\mathcal{C}_h$ and $\mathcal{C}_v$ contours. This can be understood from Eq. (6.35) where $V = 0$ for directions lying in the Π_x and Π_y planes. Moreover, the orientation of the (linear) polarization is along either the $\hat{\boldsymbol{\theta}}$ or $\hat{\boldsymbol{\phi}}$ directions since $U = 0$ from Eq. (6.34), (see Sec. 1.6 and the ellipsometric interpretation given in [2]). Figure (6.9) shows that the intensity of the wave is invariant upon reflection about the Π_x and Π_y planes which is due to the invariance of Eq. (6.30) upon reflection. The change in the wave's rotation and the orientation of its polarization ellipse for directions reflected about the Π_x and Π_y planes is accounted for by the negative sign appearing in Eqs. (6.32) and (6.33).

Comparing the rotation of the scattered wave in Fig. (6.11) to the ellipticity in Fig. (6.10) reveals that the wave becomes linearly polarized where angular regions of opposing rotation meet. In this sense the angular structure of the ellipticity and the rotation of the wave are correlated. There are also distinct points where the ellipticity of the wave becomes circular, one of which is indicated Fig. (6.10). The points of circular polarization and the lines of linear polarization appear to be examples of the the so-called "$\mathcal{C}$-points" and "$\mathcal{L}$-

lines" of [72, 73]. One can find seven such $\mathcal{L}$-lines in Fig. (6.11) and four such $\mathcal{C}$-points in Fig. (6.10).

6.1.4 Interpretation

By applying the microphysical model, the analysis in this chapter correlates the reflection symmetries of a sphere's internal field to the overall structure of the polarization state of its scattered wave. This correlation can be given a clear physical meaning by remembering that the quantities $z_{i\theta}$ and $z_{i\phi}$ represent the amplitudes of the θ and ϕ-components of each wavelet's contribution to the far-field scattered wave. These amplitudes depend both on the field at a location inside the sphere and on the direction to the observation point. In this sense, the wavelet amplitudes facilitate the connection between the internal and scattered waves.

To describe how the internal wave is connected to the scattered wave in more detail, first suppose that the observation point lies on the $\mathcal{C}_h$ contour. The sketch in Fig. (6.12) shows two wavelets located at points $\mathbf{r}_j$ and $\mathbf{r}_k$ inside a sphere. These points are mirror points about the Π_x plane and the field vectors shown at the points, $\mathbf{E}^{int}(\mathbf{r}_j)$ and $\mathbf{E}^{int}(\mathbf{r}_k)$, obey the symmetry properties of Eqs. (6.10)-(6.14). Shown at $\mathbf{r}$ are the contributions to the scattered field from these wavelets, $\mathbf{E}_j$ and $\mathbf{E}_k$, respectively. The projections of $\mathbf{E}_j$ and $\mathbf{E}_k$ onto the $\hat{\boldsymbol{\theta}}$ direction are indicated. These projections are given by the wavelet amplitudes $z_{j\theta}$ and $z_{k\theta}$, except for a factor of $\exp(ikr)/r$ which is a constant on $\mathcal{C}_h$, recall Eq. (4.30). The sketch demonstrates that these components of the wavelet's contribution to the scattered field cancel each other, and hence, it is only the ϕ-components that contribute. These components, in turn, depend only on the x-component of the internal electric field and yield a linearly polarized scattered wave on the $\mathcal{C}_h$ contour. A similar cancellation process can be used to explain the linear polarization state on the $\mathcal{C}_v$ contour.

From this cancellation-based viewpoint, one can now understand that *the ellipticity of the polarization state is due to the break in symmetry of the scattering arrangement that*

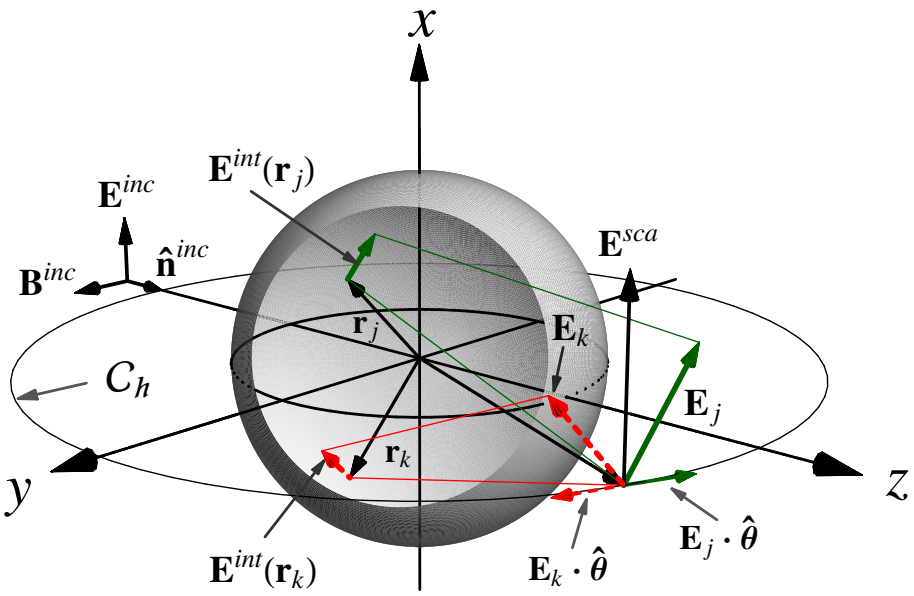

Figure 6.12: Illustration of the wavelet-component cancellation that is responsible for the linear polarization state of the scattered wave on the $\mathcal{C}_h$ contour. A similar picture applies for the $\mathcal{C}_v$ contour. Internal mirror points $\mathbf{r}_j$ and $\mathbf{r}_k$ about the Π_x plane are shown along with field vectors $\mathbf{E}^{int}(\mathbf{r}_j)$ and $\mathbf{E}^{int}(\mathbf{r}_k)$ that satisfy the symmetry conditions of Eqs. (6.10)-(6.14). Each of these internal wavelets contribute $\mathbf{E}_j$ and $\mathbf{E}_k$ to the scattered field where the cancellation of their θ-components is evident.

results from taking the observation point off the $\mathcal{C}_h$ or $\mathcal{C}_v$ contours. This break in symmetry prevents the delicate cancellation of components resulting in the linear polarization state and introduces a relative phase shift between the components. Moreover, this picture is true *regardless of the shape of the particle* provided that it is uniform and has reflection symmetry about the planes containing $\mathcal{C}_h$ and $\mathcal{C}_v$.

6.2 Spheroidal Particles

The examples considered above involve only uniform spherical particles. Consequently, the only way that the reflection symmetry of the scattering arrangement can be broken is by removing the observation point from one of the symmetry planes, Π_x or Π_y. A nonspherical particle, however, offers the opportunity to keep the observation point in the Π_x or Π_y plane and see the consequences of breaking the reflection symmetry by changing the particle orientation.

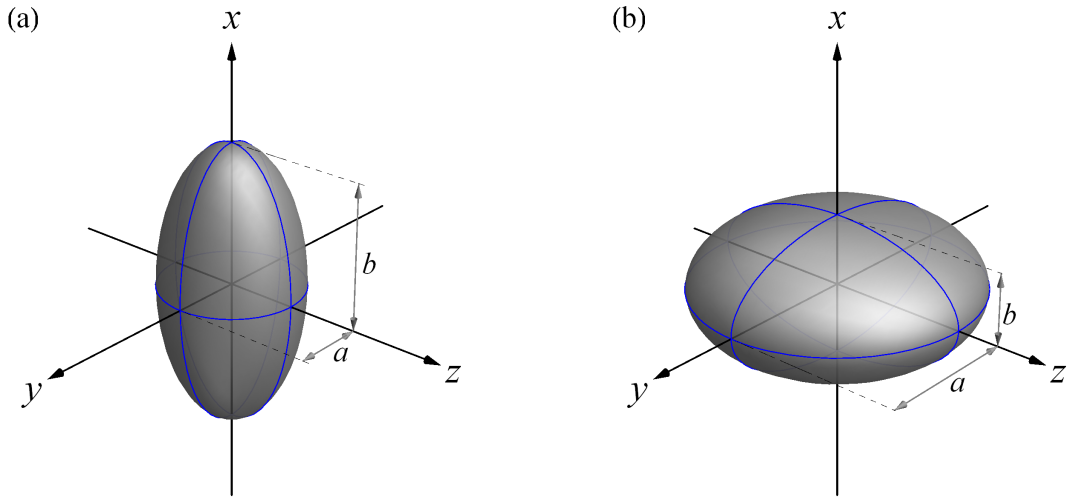

Figure 6.13: Prolate (a) and oblate (b) spheroids with minor and major axes a and b, respectively.

A simple nonspherical particle shape is the spheroid. This shape is simple because it can be described by the length of its minor and major axes, a and b, respectively, and because it has a single axis of rotational symmetry. A spheroid is prolate, or "pointy," if its aspect

ratio is $a/b < 1$, oblate, or "flat," if $a/b > 1$, and degenerates into a sphere if $a/b = 1$, see Fig. (6.13). The spheroid's volume can be compared to that of a sphere using the volume equivalent sphere radius R_{ve}, which is the radius of a sphere with the same volume.

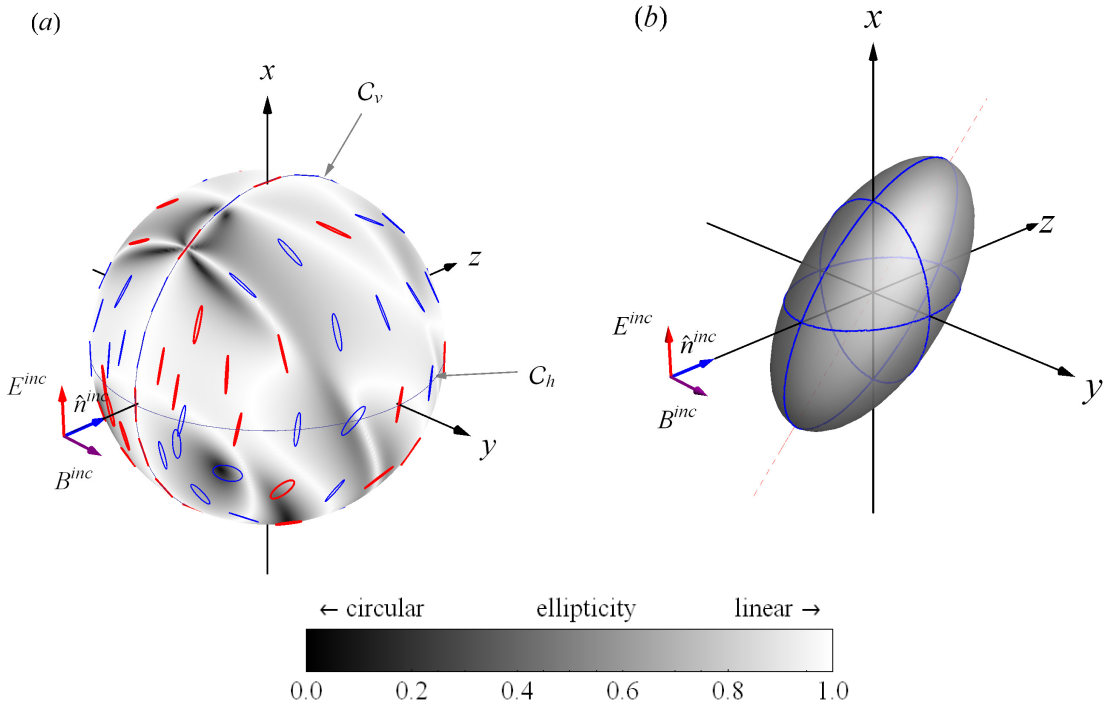

Figure 6.14: Far-field scattered wave polarization state (a) and corresponding spheroidal particle orientation (b), recall Fig. (6.8). The size, refractive index, and aspect ratio of the particle are $kR_{ve} = 3.9$, $m = 1.25 + 05i$, and $a/b = 0.5$, respectively. The rotation of the wave is shown by the same ellipse color code as in Fig. (6.8).

Plot (a) in Fig. (6.14) shows the polarization state of the far-field scattered wave due to the prolate spheroidal particle with aspect ratio $a/b = 0.5$, shown in (b). The size and refractive index of this particle are $kR_{ve} = 3.9$ and $m = 1.25 + 0i$, respectively, and the scattered wave is found with the DDA using $N = 26896$ discrete volume elements. The Euler angles of Sec. 5.4 describing the orientation of this particle are $\alpha = 0$, $\beta = \pi/4$, and $\gamma = 0$. Notice that this orientation has reflection symmetry about the Π_y plane. From the wavelet-component cancellation scheme discussed in Sec. 6.1.4, one would expect that the polarization state is linear in the Π_y plane along the $\mathcal{C}_v$ contour. Inspection of plot (a) in Fig. (6.14) shows that this is indeed the case. Now consider Fig. (6.15), with the same

prolate spheroid as in Fig. (6.14), except here the spheroid is reoriented so that there are no planes of reflection symmetry for the resulting scattering arrangement[3]. One can see from plot (a) in Fig. (6.15) that this break in symmetry removes the linear polarization along the $\mathcal{C}_v$ contour.

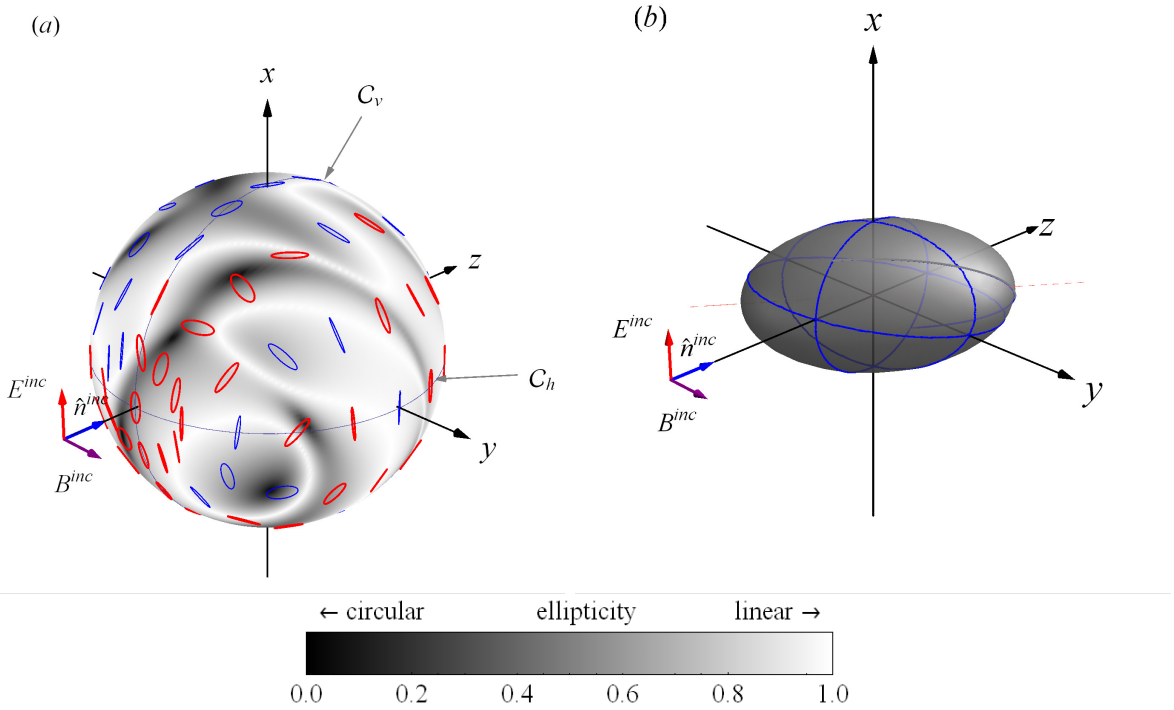

Figure 6.15: Same as in Fig. (6.14) except the particle is rotated about the z-axis to break all reflection symmetry in the scattering arrangement.

6.3 Comments

When a sphere is illuminated by a linearly polarized incident wave, as in the above, *the scattered wave is completely polarized, either linearly or elliptically, in all directions* [24, p. 780]. There is no depolarization. That does not mean, however, that the polarization state does not change, and indeed the results in this chapter remind one that it does change with angle.

[3]Recall that the scattering arrangement involves not only the particle but also the polarization and propagation directions of the incident wave.

One could, of course, derive the symmetry properties of the Stokes parameters in Eqs. (6.30)-(6.35) using the Mie *scattered* wave directly, see for example [24, 47, 63]. This wave is similar to the Mie internal wave in that it also involves an expansion in vector spherical wave functions, except with different expansion coefficients. Each term in the expansion can be regarded as a partial wave consisting of an electric and magnetic part. These partial waves essentially provide a way to describe the structure of the scattered wave in terms of component waves of simpler form. For example, the scattered wave for a small sphere ($kR \ll 1$) with a moderate, real-valued refractive index, will have an expansion that is dominated by the first electric partial wave [24]. This partial wave is equivalent in form to the far-field wave radiated by an electric dipole. For a more arbitrary sphere, higher-order partial waves will be present in the expansion. Although the partial waves are useful to describe the structure of the scattered wave, they offer no explanation for the *physical origin* of the structure; that information is encoded in a complicated way in the expansion coefficients via the boundary conditions on the surface of the particle. In using the partial waves to derive Eqs. (6.30)-(6.33), one is only able to shown that the symmetry of the Stokes parameters is due to the *mathematical* behavior of the partial waves under reflection of $\hat{\mathbf{r}}$ about the Π_x or Π_y planes. Although this is a valid mathematical account of the symmetry and has utility in its own right, it does little to provide the insight that the above microphysical approach yields.

Chapter 7

Power-Law Patterns

"I would say "q." It is q, the scattering wave vector, that is physically motivated with an inverse that describes the length scale probed by the scattering experiment."

-C. M. Sorensen [74]

The far-field scattered intensity for many particles displays a simple power-law behavior with scattering angle when represented in terms of the scattering wave vector q in so-called q-space. An empirically based study of the evolution of this power-law behavior reveals that certain features in the intensity curves, the patterns, relate to the particle's size and refractive index [75]. Consequently, the power-law patterns represent a simple inversion technique that can be used to relate the structure of measured intensity curves to certain (unknown) characteristics of a particle. The patterns then, can be quite useful in practice to study the physical properties of particulate systems *in situ*.

One can use analytical and numerical techniques, such as Mie theory and T-Matrix, to calculate the intensity curves for an arbitrary particle and see the patterns. However, the mathematical complexity of these calculations can severely hinder any attempt to understand the cause of the patterns. A simplified model, based on Huygens' principle, was proposed in [75] to explain the patterns. Although surprisingly successful in predicting aspects of the power-law patterns, the inherent assumptions involved in the model prevent it from being a convincing explanation for the cause of the patterns. In the following, the microphysical model will be used to explain the cause in terms of wavelet superposition and

in the process reveal the physical significance of the empirical relation between the patterns and the particle.

7.1 q-space

In many simple scattering measurements, the detector is located at a fixed radial distance from the particle in its far-field zone and is rotated in scattering angle from the forward direction at $\theta = 0$ to the backward direction at $\theta = \pi$ along the circular contours $\mathcal{C}_h$ or $\mathcal{C}_v$ in the horizontal or vertical scattering planes, see Figs. (6.2). The analysis in Ch. 6 shows that if the particle's shape has reflection symmetry about these planes, the angular structure of the Stokes parameters will also have reflection symmetry about these planes. For a uniform spherical particle then, one need only consider the angular dependence of the Stokes parameters along the half of the contours covering the range $0 \leqslant \theta \leqslant \pi$. This half-$\mathcal{C}_h$ contour arrangement is shown in Fig. (7.1) and will be the arrangement used for the analysis throughout this chapter.

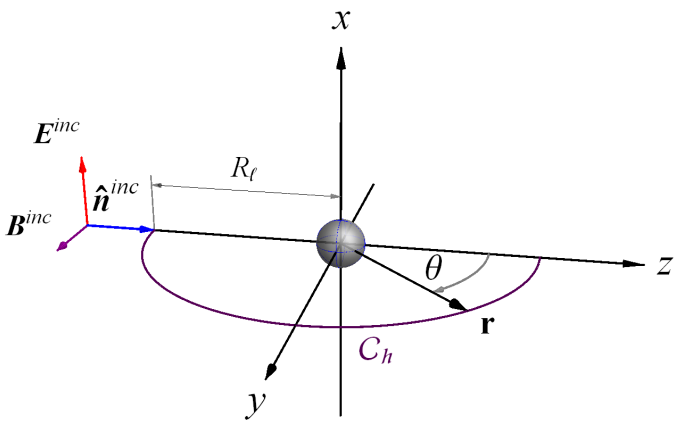

Figure 7.1: Simple scattering measurement arrangement involving a spherical particle.

The power-law patterns in the scattered intensity I are revealed when I is plotted in

terms of the scattering wave vector $\mathbf{q}$ [75]. This vector represents the difference in momentum between the incident plane wave and the far-field scattered spherical wave in the $\hat{\mathbf{r}}$ direction,

$$\mathbf{q} = k\left(\hat{\mathbf{n}}^{inc} - \hat{\mathbf{r}}\right). \tag{7.1}$$

The angular dependence of $\mathbf{q}$ is

$$q = 2k\sin\left(\frac{\theta}{2}\right). \tag{7.2}$$

Figure (7.2) shows the angular evolution of I along the half-$\mathcal{C}_h$ contour in Fig. (7.1) for a variety of spherical particles with different kR and m. The curves are calculated from Eq. (1.47) using Mie theory. Plot (a) shows the intensity curves as a function of the polar angle θ in semi-log scale, whereas (b) shows the same curves in log-log scale as a function of the dimensionless quantity qR. One can see from this latter plot that the intensity curves are roughly bounded by linear envelopes. The linear envelopes in double log scale indicate an overall power-law dependence of I on qR in addition to the detailed semi-periodic finer structure also seen. The transition from one power-law region to the next is called a crossover. In a sense, one can regard the envelopes as a coarse average of the more detailed structure. The term "q-space" refers to the analysis of scattering quantities, like the intensity, when described as a function of q rather than θ.

Now consider the phase shift parameter ρ given by

$$\rho = 2kR\,\mathrm{Re}\{m-1\}. \tag{7.3}$$

The physical meaning of the phase shift parameter is the net difference in phase between a ray passing through the center of the sphere along a diameter and a ray in the incident wave traversing the same distance in vacuum [47, p. 132], [76, p. 104]. The q-space power-law patterns show a semi-quantitative common behavior parametrized by ρ:

$$I(qR) \simeq \begin{cases} (qR)^0 & 0 \leqslant qR < 1 \\ (qR)^{-2} & 1 \lesssim qR \lesssim \rho \text{ if } \rho > 1 \\ (qR)^{-4} & \rho \lesssim qR. \end{cases} \tag{7.4}$$

Figure (7.3) shows scattering curves for very different spheres that all share the same value of ρ. One can see the common power-law behavior given by Eq. (7.4). This common behavior between such different spheres is referred to as ρ universality[1].

The power-law patterns above and several similar patterns are described in detail in [75, 77–79]. Although it is ultimately the Mie solution that yields the specific angular functionality of the scattered intensity, and hence is the origin of the power-law behavior, the mathematical complexity of the solution obscures the underlying physical significance of the patterns. The following will attempt to reveal this physical significance using the microphysical model.

[1]This is similar to the "homologous scattering diagrams" of [47, p. 183].

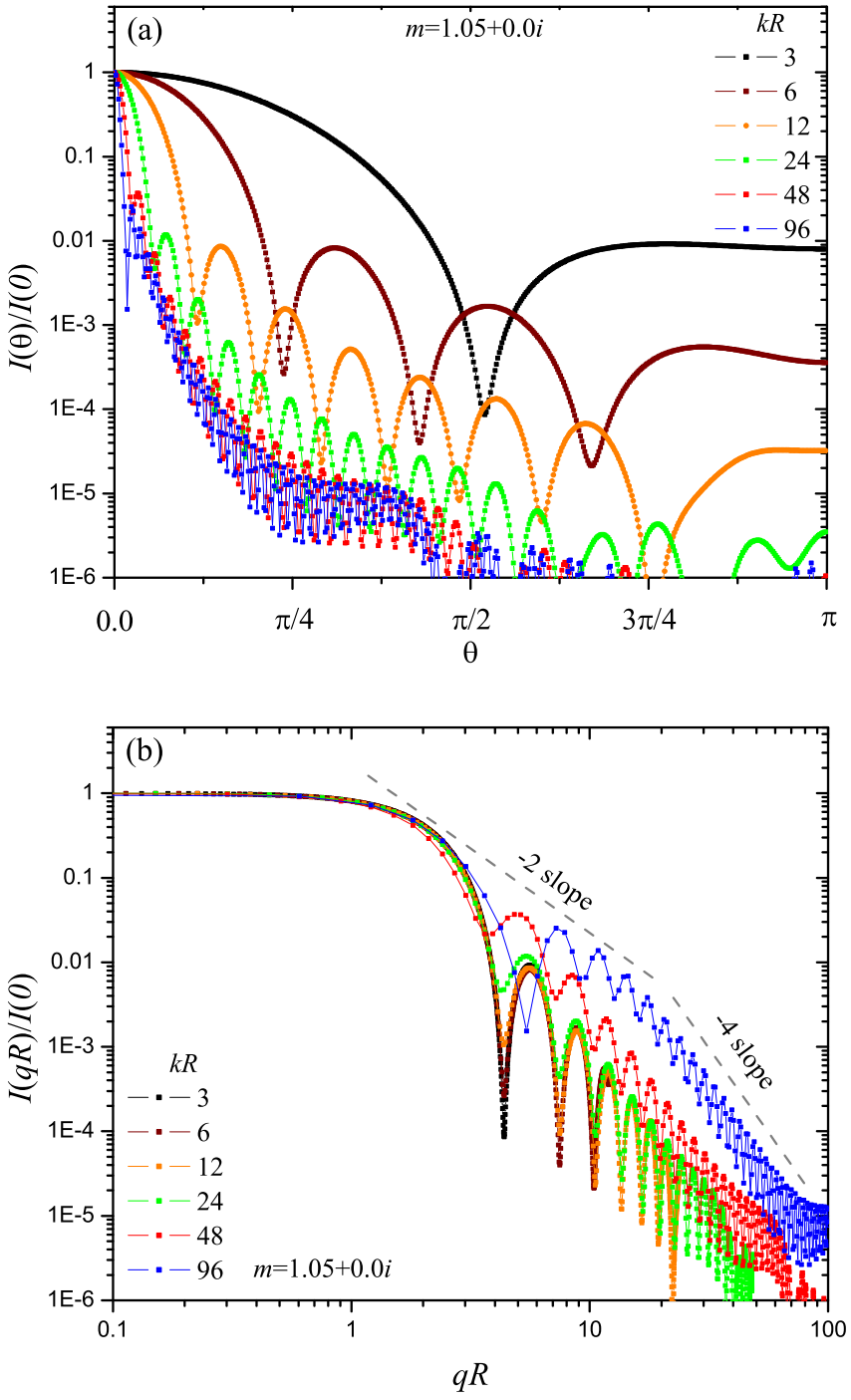

Figure 7.2: Angular power-law structure of scattered intensity for a variety spherical particles. Plot (a) shows the intensity curves plotted semi-log as a function of angle for the scattering arrangement shown in Fig. (7.1). Plot (b) shows the same curves as in (a) except plotted log-log as a function of qR. All curves are normalized to the forward direction.

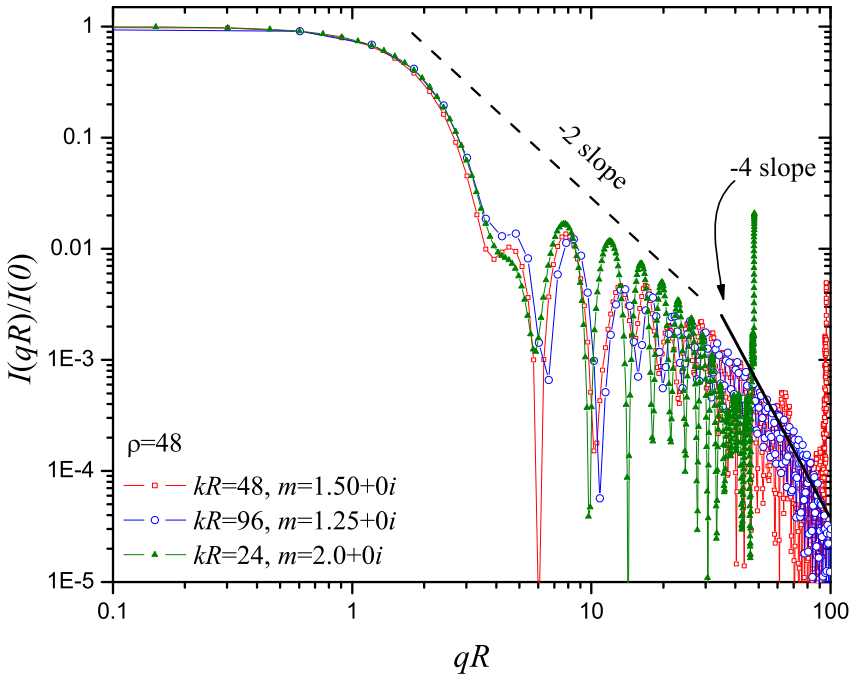

Figure 7.3: Common power-law behavior of normalized scattered intensity curves for spheres sharing the same phase shift parameter.

7.2 Small Phase-Shift Regime

Recall that the meaning of the phase shift parameter ρ is the relative change in phase between a ray passing through the center of a sphere and another ray traveling the same distance (undisturbed) outside of the sphere. This geometrical-optics based interpretation demonstrates that ρ is basically a measure of the overall degree of refraction that occurs in the sphere. If $\rho \ll 1$, refraction is weak. In the context of the microphysical model, this corresponds to weak coupling between the wavelet sources in the particle[2]. Consequently, the polarization of the particle's volume elements is dominated by the incident wave. This is the gist of the Rayleigh-Debye-Gans (RDG) approximation[3], and formally corresponds to the approximation of the particle's internal field as

$$\mathbf{E}^{int}(\mathbf{r}) \simeq \mathbf{E}^{inc}(\mathbf{r}), \quad \text{if} \quad \rho < 1. \tag{7.5}$$

One should note that the RDG approximation does not necessarily correspond to a situation that is too unrealistic to be encountered in the reality. For example, many biological cells have an average refractive index in vacuum that is only slightly different than that of water. Then, when these cells are immersed in water, as they often are in their natural environment, the *relative* refractive index of the water-cell system can be very close to unity, corresponding to $\rho < 1$ via Eq. (7.3). Another important example is when a particle is illuminated by a wave with a frequency in the X-ray range, where the vacuum refractive index of many materials is close to unity, also corresponding to small ρ.

The RDG approximation results in a substantial simplification of the VIE for observation points in the far-field zone. Recall Eq. (2.9) in combination with Eq. (2.46),

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{k^2}{4\pi}\frac{\exp(ikr)}{r}(m^2-1)\int_{V^{int}}\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}')\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathbf{r}'. \tag{7.6}$$

Since the observation point is confined to the half of the $\mathcal{C}_h$ contour in the horizontal scat-

[2]Recall the discussion of coupling between dipoles in DDA in Sec. 5.6
[3]Also called the first Born approximation.

tering plane, one can show that

$$\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}\right) = \hat{\mathbf{x}} + \hat{\mathbf{y}}, \quad \mathbf{r} \in \mathcal{C}_h. \tag{7.7}$$

Combining this with Eq. (7.6) gives

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)\int_{V^{int}}\left[E_x^{int}(\mathbf{r}')\,\hat{\mathbf{x}} + E_y^{int}(\mathbf{r}')\,\hat{\mathbf{y}}\right]\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathbf{r}', \tag{7.8}$$

where R_{en} is the radius of the $\mathcal{C}_h$ contour. Now recall the wavelet-component cancellation arguments of Sec. 6.1.4. The reflection symmetry of this scattering arrangement shows that contributions to the integral in Eq. (7.8) from the y-component of the internal field cancel each other. Then, they can be ignored, giving

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)\int_{V^{int}}E_x^{int}(\mathbf{r}')\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathbf{r}'\,\hat{\mathbf{x}}. \tag{7.9}$$

Now the RDG approximation is made by combining this with Eqs. (3.1) and (7.9) to give

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)E_o^{inc}\int_{V^{int}}\exp\left[ik\left(\hat{\mathbf{n}}^{inc}-\hat{\mathbf{r}}\right)\cdot\mathbf{r}'\right]\mathrm{d}\mathbf{r}'\,\hat{\mathbf{x}}. \tag{7.10}$$

Notice however, that the argument of the exponential in this integral is equivalent to the scattering wave vector $\mathbf{q}$ of Eq. (7.1), then

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)E_o^{inc}\int_{V^{int}}\exp(i\mathbf{q}\cdot\mathbf{r}')\,\mathrm{d}\mathbf{r}'\,\hat{\mathbf{x}}. \tag{7.11}$$

Comparison of this result with Eq. (1.16) shows that *the far-field scattered wave in the RDG approximation for* $\mathbf{r} \in \mathcal{C}_h$ *is given by the Fourier transform of the the particle volume into q-space.* This will be seen again in Ch. 9.

To apply the microphysical model to the RDG sphere, the sphere volume is, as usual, divided into elements ΔV, thus discretizing the integral in Eq. (7.11),

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)E_o^{inc}\sum_i\exp(i\mathbf{q}\cdot\mathbf{r}_i)\,\Delta V\,\hat{\mathbf{x}}.$$

From this, one can define a phasor sum like Eq. (4.30) as

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{\exp(ikR_{en})}{R_{en}}\sum_i z_i(\hat{\mathbf{r}})\,\hat{\mathbf{x}}. \tag{7.12}$$

where the phasors z_i are

$$z_i(\hat{\mathbf{r}}) = \frac{k^2}{4\pi}(m^2-1)E_o^{inc}\exp(i\mathbf{q}\cdot\mathbf{r}_i)\,\Delta V. \tag{7.13}$$

Notice that these phasors are equivalent to the $z_{i\phi}$ phasors of Eq. (4.30) assuming the RDG approximation.

From Eq. (7.12) one can define a scattering amplitude via Eq. (4.17) as

$$\mathbf{E}_1^{sca}(\hat{\mathbf{r}}) = \sum_i z_i(\hat{\mathbf{r}})\,\hat{\mathbf{x}}. \tag{7.14}$$

Then, in analogy to Eqs. (1.47) and (6.26), and since $\mathbf{r}\in\mathcal{C}_h$, the products $\hat{\boldsymbol{\theta}}\cdot\mathbf{E}_1^{sca}=0$ and $\hat{\boldsymbol{\phi}}\cdot\mathbf{E}_1^{sca}=\hat{\mathbf{x}}\cdot\mathbf{E}_1^{sca}$ so that

$$I(\hat{\mathbf{r}}) = I(q) = \frac{1}{2}\sqrt{\frac{\varepsilon_o}{\mu_o}}\left|\sum_i z_i(q)\right|^2 \quad \mathbf{r}\in\mathcal{C}_h, \tag{7.15}$$

where Eq. (7.2) is used to substitute q for scattering angle θ.

The phase of the wavelet-phasors z_i is determined by Eq. (7.13). As the observation point moves along the half-contour $\mathcal{C}_h$ from $\theta=0$ to π, the vector $\mathbf{q}$ changes and hence the phase of z_i also changes. One can see from Eq. (7.13) that the magnitude of all the phasors is the same, which is due to the uniform magnitude of the incident wave and thus the phasors are spread out in the complex plane in a ring. The radius of the ring is the common magnitude of the phasors and the location of a given phasor on the ring is determined by the scattering-angle dependent product $\mathbf{q}\cdot\mathbf{r}_i$ in the phase of Eq. (7.13).

To study the wavelet superposition in more detail, consider how the phasors are distributed in the complex plane when $q=0$, corresponding to the forward direction ($\theta=0$). In this case, the exponential in Eq. (7.13) is unity and all the phasors are in phase. The resulting scattered intensity is then

$$I(0) = \sqrt{\frac{\varepsilon_o}{\mu_o}}\frac{k^4}{32\pi^2}|m^2-1|^2\left(E_o^{inc}\right)^2\left(V^{int}\right)^2, \tag{7.16}$$

where use has been made of the relation $\sum_i \Delta V = V^{int}$. This result shows that in the RDG approximation a sphere's forward scattered intensity is proportional to the square of

its volume V^{int}, and this happens because all of the sphere's wavelets are in phase, i.e., constructively interfere.

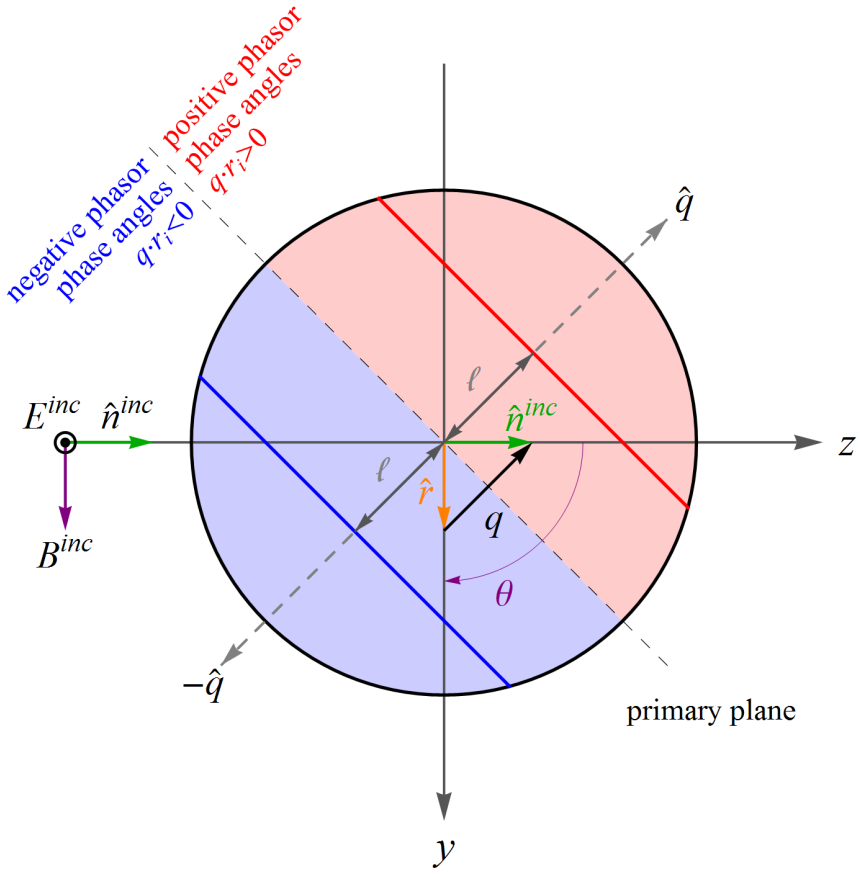

Figure 7.4: Horizontal scattering-plane slice through a spherical particle. The direction of the scattering wave vector $\mathbf{q}$ is shown for a scattering angle of $\theta = \pi/2$. The hemisphere shaded red (blue) denotes the location of wavelets with phasors having positive (negative) phase angles.

Returning to Eq. (7.1), one can see that as θ increases from zero, the projection of $\mathbf{q}$ onto a wavelet's source location $\mathbf{r}_i$ will deviate from zero and cause the phase in the exponential of Eq. (7.13) to change. Notice that, because $\mathbf{r} \in \mathcal{C}_h$, $\mathbf{q}$ lies completely in the horizontal scattering plane. Now consider Fig. (7.4), which shows the scattering plane slice through the sphere. Three planes are shown; a primary plane and two parallel planes. The primary plane passes through the sphere center and is oriented with $\hat{\mathbf{q}}$ as its normal. The other two planes are parallel to the primary plane. All three planes are perpendicular to the horizontal

scattering plane. One of the two parallel planes is displaced by a distance l from the primary plane along the $\hat{\mathbf{q}}$ direction, and the other plane is displaced by the same distance along the $-\hat{\mathbf{q}}$ direction. The locations $\mathbf{r}_i$ contained in these displaced planes have equal magnitude but projections along $\mathbf{q}$ that are opposite in sign. The phase of the phasors corresponding to locations in the plane displaced along $\hat{\mathbf{q}}$ are positive since $\mathbf{q} \cdot \mathbf{r}_i > 0$, see the red shaded hemisphere in Fig. (7.4). In contrast, the phase angle of the phasors associated with the plane displaced along $-\hat{\mathbf{q}}$ have the same magnitude but opposite sign, since $\mathbf{q} \cdot \mathbf{r}_i < 0$, see the blue shaded hemisphere in Fig. (7.4).

The significance of this phase angle behavior is that phasors belonging to the hemisphere containing $\mathbf{q} \cdot \mathbf{r}_i > 0$ rotate counter-clockwise in the complex plane as θ increases from zero, whereas the phasors in the other hemisphere, $\mathbf{q} \cdot \mathbf{r}_i < 0$, rotate clockwise. Consequently, *cancellation* between phasors belonging to these hemispheres begins to occur under the summation of Eq. (7.14). Recall from Fig. (4.3) that two phasors opposed by π in phase cancel when added together. The cancellation of phasors belonging to the two hemispheres means that the scattered intensity decreases with advancing q, (or equivalently, with θ), as compared to its value in the forward direction.

Fig. (7.5) shows the normalized q-space scattered intensity curve as calculated from Mie theory for a spherical particle with $kR = 4.0$ and $m = 1.05 + 0i$. This corresponds to a phase shift parameter of $\rho = 0.4$, placing this sphere within the range of validity of the RDG approximation. Four values of qR are indicated on the curve and the phasor distributions for each of these qR values are shown in Fig. (7.6). By comparing these phasor plots to the respective points on the scattering curve, one can see how the phasors spread out in a counter-rotating manner to form a ring-like distribution[4]. Note that the radial spread of the phasors, which give the ring-like distribution a "thickness," is due to the departure of the exact (Mie) internal field from the RDG approximation.

The next task is to show how phasor cancellation accounts for the particular power-law

[4]The clumped appearance of the phasors is probably due to the finite discretization of the sphere volume.

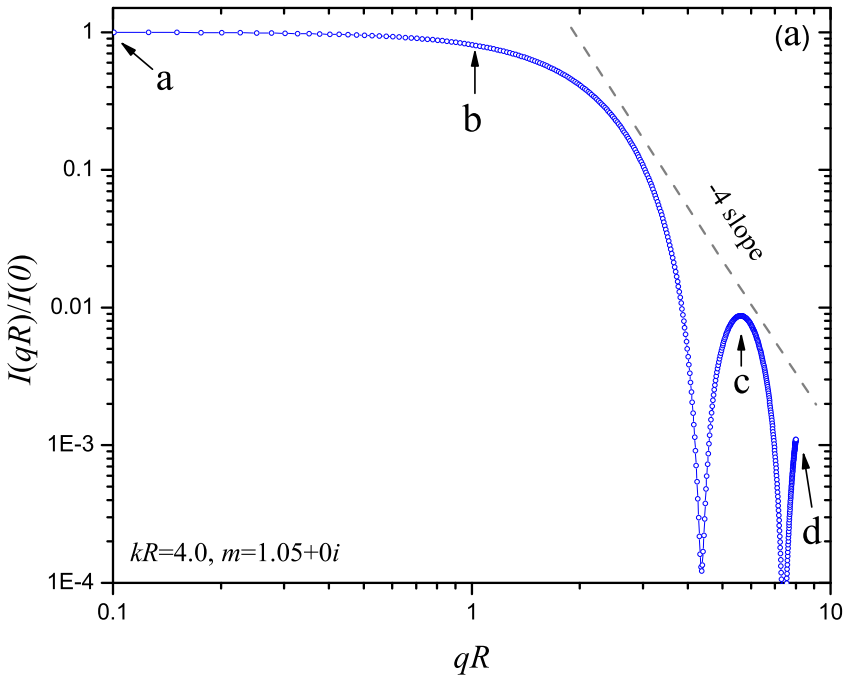

Figure 7.5: Normalized scattered intensity curve for a spherical particle with $kR = 4.0$ and $m = 1.05 + 0i$. The four points labeled $a - d$ correspond to the phasor plots in Fig. (7.6).

dependence of I on qR given in Eq. (7.4) for $\rho < 1$. Recall from Fig. (7.4) and Eq. (7.13) that all phasors corresponding to volume elements centered on a plane with normal $\hat{\mathbf{q}}$ share the same phase angle; the angle depends only on the displacement of these (parallel) planes from the primary plane. Then, consider dividing the sphere's volume into a total of N slices formed by a sequence of these planes separated by l, see Fig. (7.7). The thickness l of a slice is determined by how much the phasor phase-angle changes across the slice. Notice that the phase angle of all phasors associated with the primary plane is zero because every $\mathbf{r}_i$ in that plane is by definition perpendicular to $\mathbf{q}$. Then, the distance from the primary plane to the very next plane is given by the requirement that the change of phase between the planes is $\pm\pi$. Remember that the angle is either positive or negative depending on which side of the primary plane $\mathbf{r}_i$ lies. Then the slice thickness is given by

$$ql = \pi, \tag{7.17}$$

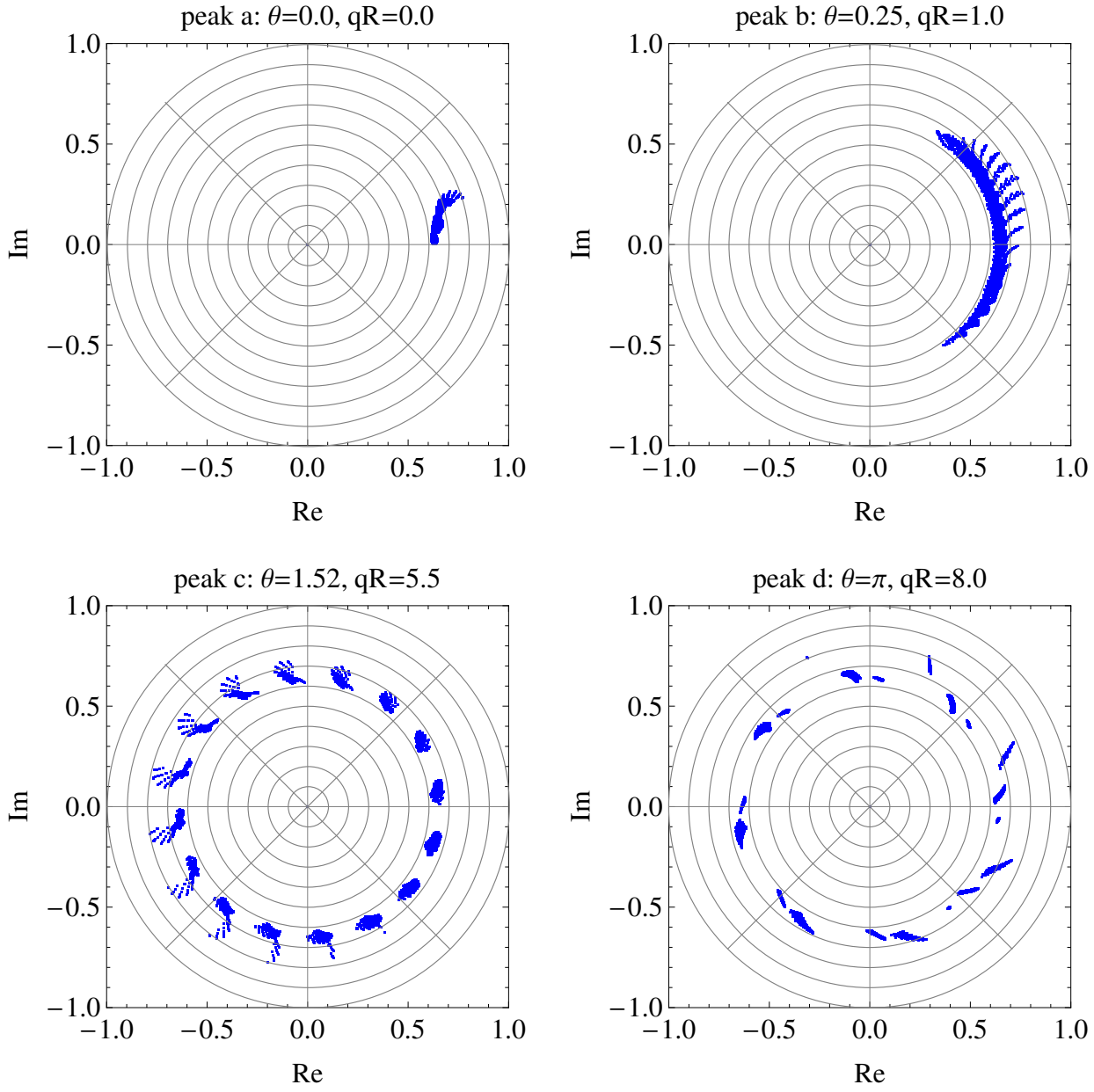

Figure 7.6: Phasor plots corresponding to the points $a-d$ indicated on the scattering curve in Fig. (7.5). Each dot represents the phasor belonging to a wavelet in the particle. The scattering angles and qR values are indicated. The magnitude of the phasor distributions have been globally scaled to a value near one for appearance sake; the important aspect of these distributions, however, is their angular distribution.

which shows that as θ increases from zero, and thus increases q, the slice thickness decreases.

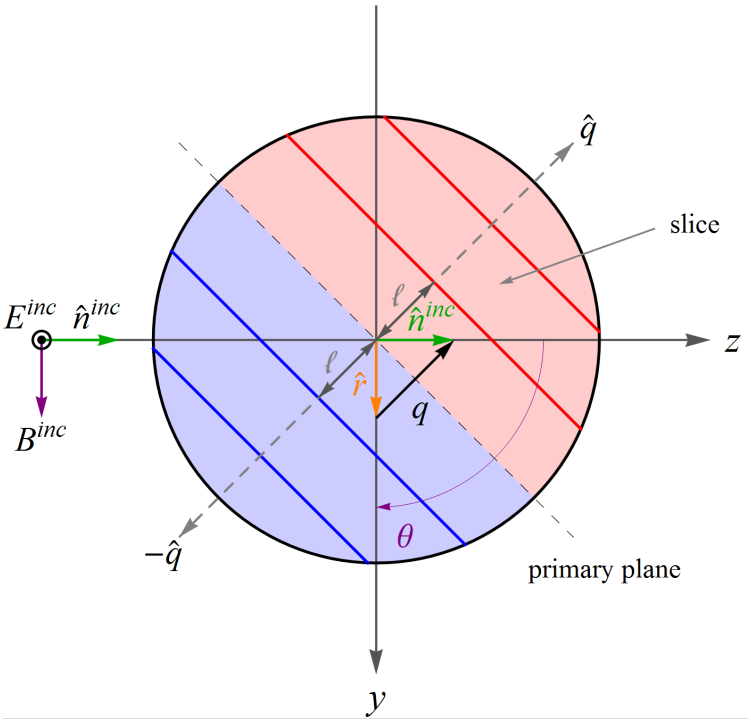

Figure 7.7: Like Fig. (7.4), but showing a sphere's slice.

In the forward direction, where $\theta = 0$, Eqs. (7.2) and (7.17) show that the slice thickness is infinite. There are only two slices that encompass the sphere volume in this case; one slice for phasors with positive phase angles and the other slice for negative phase angles. As the θ increases, there will be a point where the slice thickness decreases enough that the phase angle of the phasors in the "caps" of the sphere along the $\hat{\mathbf{q}}$ and $-\hat{\mathbf{q}}$ directions will reach $\sim \pi/2$ and $\sim -\pi/2$, respectively. Let the value of q corresponding to the scattering angle at which this occurs be denoted q_1. At this scattering angle, these cap-phasors cancel out each other's contribution to the scattering amplitude in Eq. (7.14); *the wavelets associated with these phasors destructively interfere, thus diminishing the intensity of the far-field scattered wave*. The location of the tip-phasors is at $\mathbf{r}_1 = R\hat{\mathbf{q}}$ and $\mathbf{r}_2 = -R\hat{\mathbf{q}}$, and hence q_1 is given

by

$$\begin{aligned}\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2) &= 2q_1R,\\ &= \pi,\end{aligned}$$

or equivalently,

$$q_1R \simeq \frac{\pi}{2}, \quad \text{or} \quad q_1 \sim \frac{1}{R}. \tag{7.18}$$

Equation (7.18) identifies the first crossover, $(qR)^0$ to $(qR)^{-2}$ or $(qR)^{-4}$, in the log-log q-space intensity plots shown plot (b) in Fig. (7.2). This crossover is referred to as the Guinier crossover due to its extensive use X-ray scattering [77, 80]. The Guinier crossover occurs for any particle shape; if it is nonspherical, the radius of gyration R_g usually takes the place[5] of R in Eq. (7.18) [80, Sec. 2.1.3]. The utility of the Guinier crossover in practice is that it can provide an estimate of the particle size without knowledge of the particle's refractive index.

Now return to Fig. (7.7). As the scattering angle increases beyond the Guinier crossover, the thickness of the slices decreases via Eq. (7.17), while the total number of slices increases. Because the change in phasor phase-angle across the width of a single slice is π, the phasors associated with adjacent slices partially cancel each other. The cancellation is not complete because the respective volumes of adjacent slices are not equal due to the curvature of the sphere surface. Now approximate the volume of the the n^{th} slice by the disk-shaped volume V_n with a rectangular cross section that accommodates all of the true slice's volume with a little extra, see sketch (a) in Fig. (7.8). The location of V_n is at a distance nl from the primary plane, where n is a positive or negative integer depending on which side of the primary plane the slice resides. Explicitly, the volume of the n^{th} rectangular disk-shaped slice is

$$V_n = \pi l\left(R^2 - n^2l^2\right). \tag{7.19}$$

[5]This is really true only because one usually considers a nonspherical particle in an ensemble of random orientations and hence is interested in the orientationally averaged particle size.

The degree to which the phasors associated with two adjacent slices cancel each other is given by the difference of their volumes. In terms of the disk-shaped approximations for these slices, the difference between the n^{th} and $(n+1)^{\text{th}}$ slice, denoted ΔV_n, is

$$\begin{aligned}\Delta V_n &= V_n - V_{n+1}, \\ &= \pi l \left(R^2 - n^2 l^2\right) - \pi l \left(R^2 - n^2 l^2 - 2nl^2 - l^2\right), \\ &= \pi l^3 \left(2n - 1\right). \end{aligned} \tag{7.20}$$

The pair-wise cancellation of adjacent slices continues for all n (positive and negative) moving out from the primary plane until the caps of the sphere are reached. Here, the pairing stops, resulting a spherical-cap volume. Two such caps are formed, one at each tip of the sphere. Since these caps are at the same distance from the primary plane, the phasors associated with them do not cancel, but rather add constructively in their contribution to the sum in Eq. (7.14). Now approximate the cap volume by the disk-like volume of rectangular cross section with the same base radius and thickness as the cap, and call this volume V_{cap}. Notice that the thickness of V_{cap} is not necessarily l like it is for the slices. This is because n is an integer and the sphere radius R is not, in general, an integer multiple of l. Let the thickness of the caps be s and the cap base radius be x, then from inspection of sketch (a) in Fig. (7.8), one finds that

$$\begin{aligned} s &= R - \sqrt{R^2 - x^2} \\ &= R - R\sqrt{1 - \frac{x^2}{R^2}} \\ &\simeq R - R\left(1 - \frac{x^2}{2R^2}\right) \\ &\simeq \frac{x^2}{2R}. \end{aligned}$$

Since the interest is ultimately only in the average[6] behavior of the phasor cancellation for q as it varies over its largest range, the cap thickness s can be approximated by its average

[6]Recall "average" corresponds to the linear envelopes of the power-law patterns.

value of $l/2$. Then

$$V_{cap} \simeq \frac{\pi}{2} l^2 R, \tag{7.21}$$

where the square root in the relation between s and x has been approximated since $x/R \ll 1$ for large q.

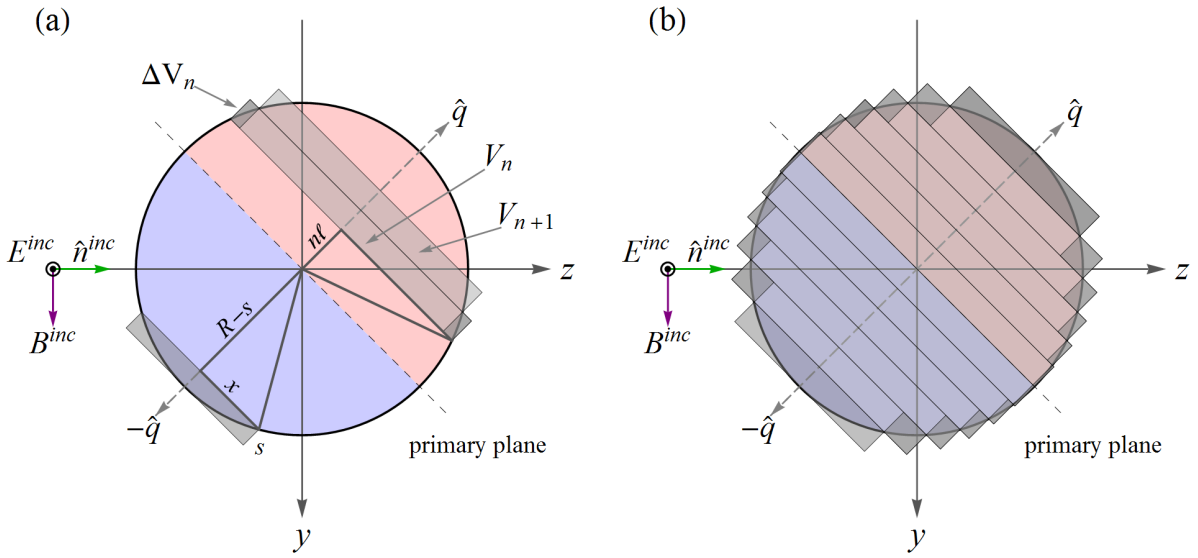

Figure 7.8: Slice-like and end-cap volumes. Sketch (a) shows the geometry used to derive the Porod scattering law in Eq. (7.24). Here the darker shaded regions are ΔV_n and represent the phasors associated with adjacent slices that do not cancel each other's contribution to the scattered intensity. Sketch (b) shows how the volumes ΔV_n change at different locations near the sphere surface.

The cancellation of adjacent slices means that the portion of the sphere's volume that is divisible into complete slices contributes to the sum in Eq. (7.14) an amount that is approximately given by the sum of the "left-over" volumes ΔV_n. Additionally, both caps contribute to the sum. Then, the average value of the scattered intensity given by Eq. (7.15) is

$$I(q) \simeq \sqrt{\frac{\varepsilon_o}{\mu_o}} \frac{k^4}{32\pi^2} |m^2 - 1|^2 \left(E_o^{inc}\right)^2 \left(2V_{cap} + \sum_n \Delta V_n\right)^2. \tag{7.22}$$

From Fig. (7.8), one can see that as q increases and l decreases, the leftover volumes ΔV_n and the end-cap volume V_{cap} decrease. However, they do not decrease at the same rate (with

increasing q). To see this consider the ratio of the volumes in the limit that $l \to 0$

$$\begin{aligned}\lim_{l\to 0}\frac{\Delta V_n}{V_{cap}} &= \lim_{l\to 0}\frac{\pi l^3\,(2n-1)}{\frac{\pi}{2}l^2 R}\\ &= \lim_{l\to 0}\frac{2l(2n-1)}{R}\\ &= 0. \end{aligned} \tag{7.23}$$

This shows that as q increases, the contribution to the scattered intensity coming from V_{cap} dominates the contribution from the leftover volumes ΔV_n. Consequently, using Eqs. (7.21) and $V^{int} = (4\pi/3)R^3$, the normalized scattered intensity can be approximated as

$$\begin{aligned}\frac{I(q)}{I(0)} &\simeq \left(\frac{V_{cap}}{V^{int}}\right)^2\\ &\simeq \left(\frac{l}{R}\right)^4,\end{aligned}$$

which from Eq. (7.17) becomes

$$\frac{I(qR)}{I(0)} \sim \frac{1}{(qR)^4}, \tag{7.24}$$

which agrees with Eq. (7.4). This result shows that the power-law behavior of the scattered intensity for a RDG sphere ($\rho < 1$) is due to the incomplete phasor cancellation associated with the end-cap volumes. Since the average evolution of the size of these cap volumes is ultimately controlled by the surface of the sphere, one can understand that this large q regime of the intensity curves "senses," or "probes" the particle surface. In the X-ray scattering community, this large q behavior of the scattered intensity is referred to as Porod's law [80, Sec. 2.1.2.4]. Porod's law is typically derived from an autocorrelation based approach that is far less illuminating than the above in the sense that one does not as easily see the connection to the particle surface.

The considerations thus far are not necessarily restricted to spherical particles and can be applied to nonspherical particles provided that the RDG approximation remains valid. The essential conclusion from the above is that the large q power-law behavior of the scattering curve is due to the scaling of the size of the end-cap volumes V_{cap}. It is relatively straightforward then, to apply this same calculation to a spheroidal particle.

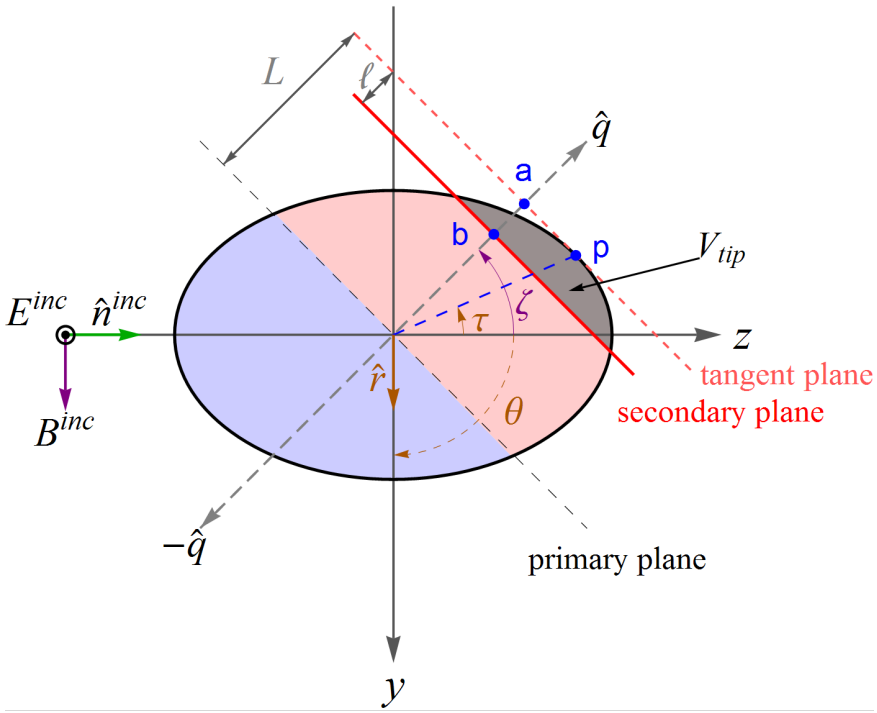

Figure 7.9: Tip volume V_{tip} (dark shade) of a prolate spheroid and the various quantities used in the text to calculate this volume. The scattering angle θ is the same as in Fig. (7.7), but the slice-like volumes shown in that figure are omitted here.

Consider the sketch in Fig. (7.9), which is the analog to sketch (a) of Fig. (7.7) for a prolate spheroidal particle. The rotational axis of the spheroid coincides with the z-axis. Recall from Fig. (6.13) that the spheroid is prolate or oblate depending on the ratio of its minor and major axes a and b, respectively. The analog to the end-cap volume of the spherical particle is shown in dark gray shade and is called the tip volume V_{tip} to avoid confusion with V_{cap}. As mentioned, the scaling of the size of the two tip volumes will establish the large q behavior of the scattering curve. The tip volume is defined by the volume contained between the spheroid surface and the secondary plane shown in Fig. (7.9). The secondary plane is parallel to the primary plane and oriented normal to $\hat{\mathbf{q}}$. Parallel to the secondary plane is the tangent plane. This plane contacts the spheroid's surface at only a single point, called the tangent point, and is labeled `p` in the sketch. The separation between this plane and the primary plane, denoted by L, will vary with $\hat{\mathbf{q}}$ due to the spheroid's shape. The thickness of the tip volume, labeled l, is determined by the separation between the tangent and secondary planes. Following the same phasor-cancellation based arguments as above, this thickness will scale inversely with q as given by Eq. (7.17).

One can see from Fig. (7.9) that the nonspherical shape substantially complicates an attempt to find an analytical expression for the tip volumes as a function of θ. Consequently, this volume is calculated numerically. It is possible, however, to find the analytical expression for the surfaces bounding the tip volume. The surface of the spheroid is given by

$$\frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} = 1. \tag{7.25}$$

The equation for the secondary plane can be found as follows. The components of the tangent point `p` in Fig. (7.9) in terms of the angle τ are

$$y_p = -a \sin\tau, \quad z_p = b \cos\tau, \tag{7.26}$$

and the equation for the primary and tangent planes are

$$y = z \tan\left(\frac{\theta}{2}\right), \tag{7.27}$$

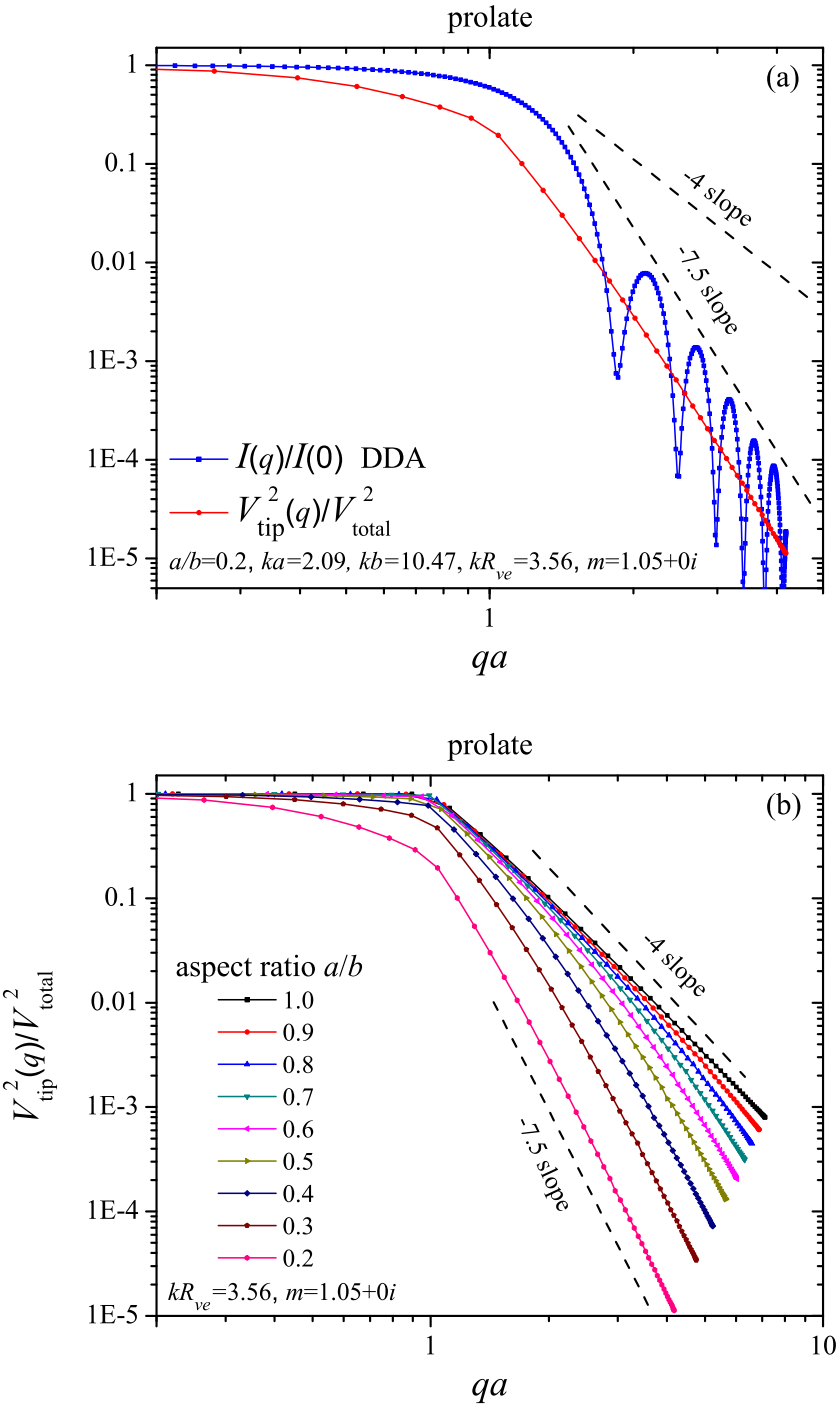

Figure 7.10: Plots of the scattered intensity corresponding to prolate spheroids. Plot (a) shows the intensity as calculated from the DDA and from the tip volumes V_{tip}. Plot (b) shows the tip-volume calculated curves for prolate spheroids with the same kR_{ve} but different aspect ratios.

$$y = \tan\left(\frac{\theta}{2}\right)(z - z_p) + y_p, \tag{7.28}$$

where use has been made of the equivalent slope $\tan(\theta/2)$ of the planes. The components of the point a can be found by realizing that the line connecting point a intersects the tangent plane at a. This gives,

$$y_a = -\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\left[y_p \sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right) + z_p \cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\right], \tag{7.29}$$

$$z_a = \cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\left[y_p \sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right) + z_p \cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\right], \tag{7.30}$$

from which one has the length of the segment L as

$$L = \sqrt{y_a^2 + z_a^2}. \tag{7.31}$$

Now the components of point b are

$$y_b = -(L - l)\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right), \quad z_b = (L - l)\cos\left(\frac{\pi}{2} - \frac{\theta}{2}\right), \tag{7.32}$$

and finally, one has the equation for the secondary plane

$$y = \tan\left(\frac{\theta}{2}\right)(z - z_b) + y_b. \tag{7.33}$$

From Eqs. (7.25) and (7.33), one can numerically calculate the tip volume V_{tip} as a function of θ (or q). Then one can find the scattered intensity from Eq. (7.22), where the contribution from the smaller volumes ΔV_n are neglected for the same reasons as with spherical particles.

Plot (a) in Fig. (7.10) shows the scattered intensity curve for a prolate spheroid as calculated from the DDA and from Eq. (7.22) using the tip volumes V_{tip}. The aspect ratio of the spheroid is needle-like $a/b = 0.2$, with a sphere volume equivalent size parameter of $kR_{ve} = 3.56$ and refractive index of $m = 1.05 + 0i$. These values are well within the range of validity of the RDG approximation. The particle's orientation is the same as shown in Fig. (7.9). One can see that the DDA curve shows the expected sequence of maxima and minima with an overall power-law behavior like the spheres above. Notice though that the slope is -7.5, unlike -4 for spheres. Comparison of this DDA generated curve to the tip

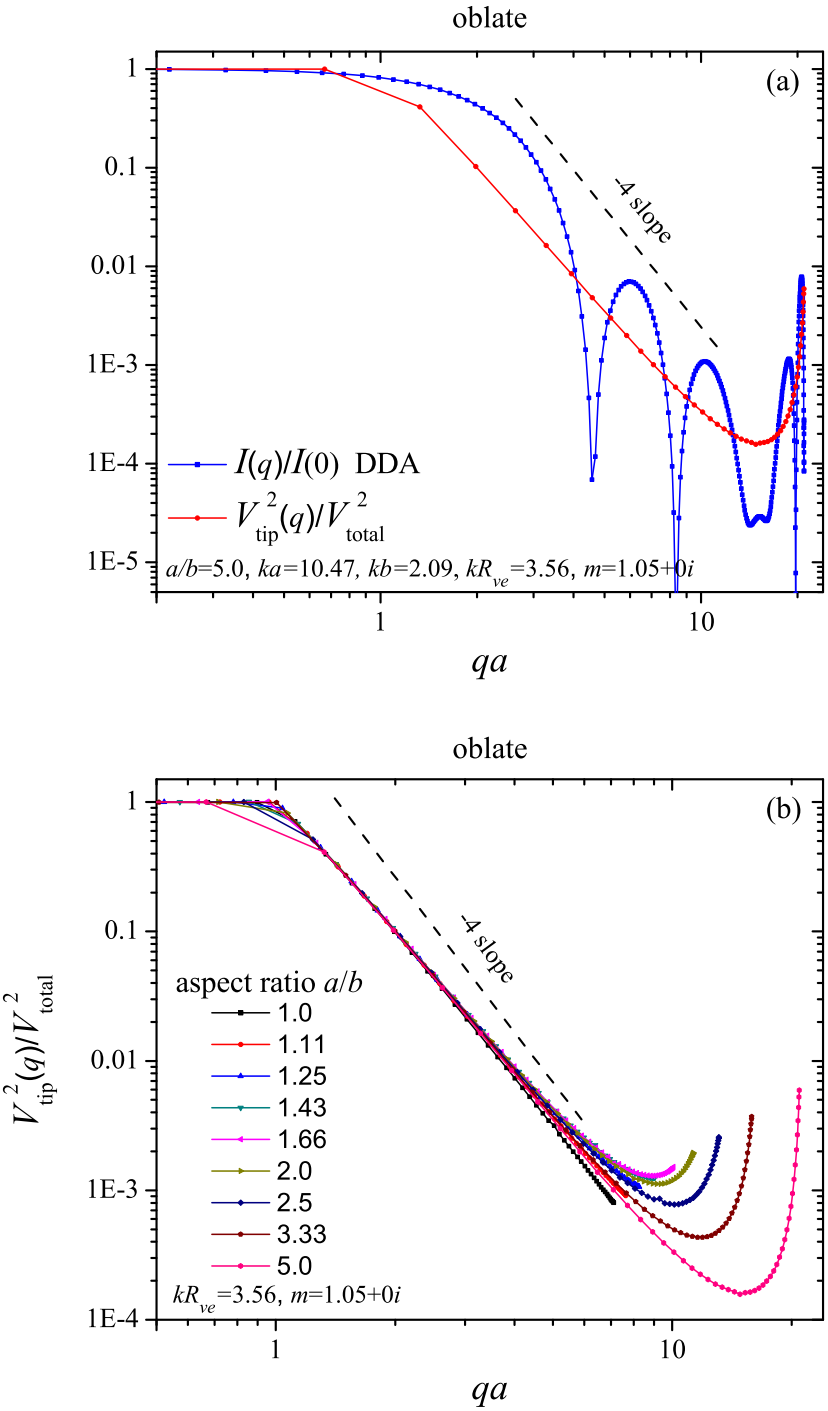

Figure 7.11: Same as Fig. (7.10) except for oblate spheroids.

volume generated approximation shows that the power-law behavior is well explained by the scaling of the tip volume with q. Hence, *exactly the same phasor-cancellation based picture developed for the spherical particle applies to this prolate spheroid.* Plot (b) in this figure shows the tip-volume generated curves for an assortment of prolate spheroids with the same kR_{ve} as in (a), but with different aspect ratios. When the aspect ratio is $a/b = 1$, the spheroid degenerates into a sphere and one can see that the power-law slope is -4 as expected. As the spheroid becomes more pointed, the slope evolves to -7.5. Careful study of Fig. (7.9) shows that the cause of this enhanced decay in the scattered intensity is due to the change in the curvature of the spheroid's surface with the decreasing eccentricity.

Figure (7.11) shows the same type of plots as in Fig. (7.10) except for oblate spheroids. Here one can see excellent qualitative agreement between the DDA and tip-volume generated curves. The up-turn in the curves in the large q range for the flatter spheroids is due to specular-like backscattering. One can understand this by realizing that the flatter the oblate spheroid gets, i.e. the larger a/b is, the more the particle appears like a small disk that is oriented perpendicular to the propagation direction of the incident wave. In contrast, when $a/b = 1$ one again sees that the -4 slope is present, as expected, since the spheroid degenerates into a sphere in this case.

An important conclusion from this analysis is that the Porod scattering law does not necessarily associate a power-law slope of -4 with the surface of a particle. Indeed, the spheroids above show a range of slopes, and one will see in Ch. 8 that non-integer slopes are typical for more complicated particle shapes.

7.3 Large Phase-Shift Regime

When the phase shift parameter of a spherical particle exceeds unity, the RDG approximation of Eq. (7.5) becomes invalid. This happens because the coupling between the volume elements in the sphere is no longer weaker than the influence of the incident wave. This means that in the context of the phasor cancellation scheme above, is that it is no longer

possible to determine how the wavelet phasors cancel based on the evolution of the slice, cap, and tip volumes with q. Nevertheless, by studying how the phasors redistribute themselves in the complex plane as q varies, one will see a second crossover at $qR \sim \rho$ appear and will also see the physical cause for it.

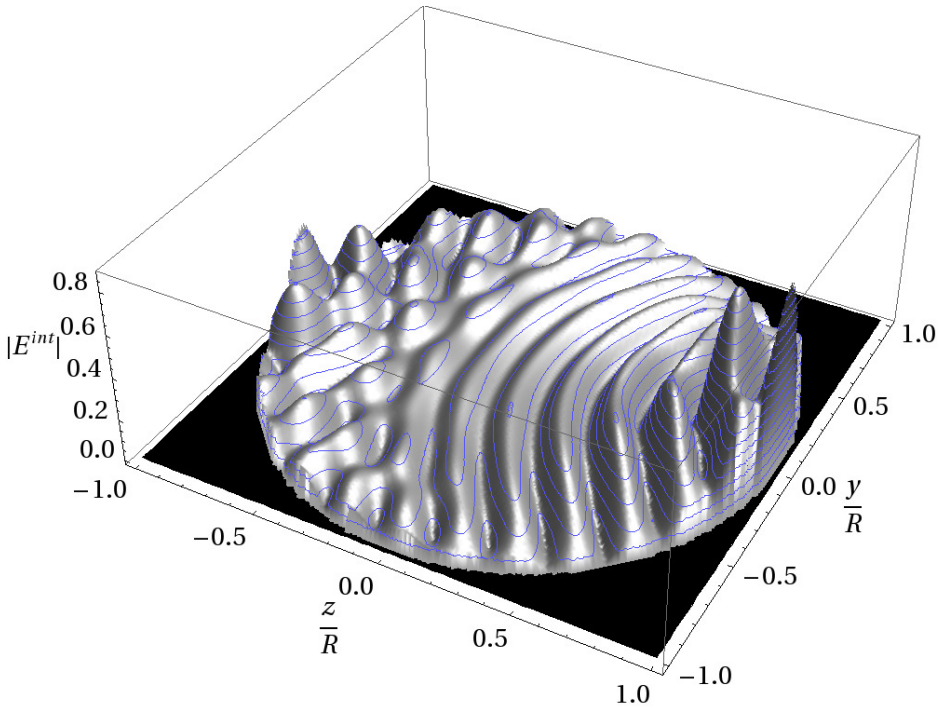

Figure 7.12: Normalized magnitude of the internal electric field in the horizontal scattering plane slice through a sphere. The size and refractive index are $kR = 12$ and $m = 1.5 + 0i$, respectively, corresponding to $\rho = 12$. The incident wave here is polarized along the x-axis and propagates from left to right along the z-axis.

Since the RDG approximation cannot be made when $\rho > 1$, Mie theory must be used to generate the sphere's internal field. The phasors that describe the scattered intensity at $\mathbf{r}$ as shown in Fig. (7.1) can be identified from Eq. (7.9),

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)\int_{V^{int}} E_x^{int}(\mathbf{r}')\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathbf{r}'\,\hat{\mathbf{x}}.$$

Discretization of the integral over the sphere volume gives

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\frac{\exp(ikR_{en})}{R_{en}}(m^2-1)\sum_i E_x^{int}(\mathbf{r}_i)\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i)\,\Delta V\,\hat{\mathbf{x}}.$$

The wavelet phasor sum follows as

$$\mathbf{E}^{sca}(R_{en}\hat{\mathbf{r}}) = \frac{\exp(ikR_{en})}{R_{en}} \sum_i z_i(\hat{\mathbf{r}})\, \hat{\mathbf{x}}. \tag{7.34}$$

which is identical to Eq. (7.12) except now the phasors are given by

$$z_i(\hat{\mathbf{r}}) = \frac{k^2}{4\pi}(m^2 - 1) E_x^{int}(\mathbf{r}_i) \exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}_i)\, \Delta V. \tag{7.35}$$

In the RDG analysis in Sec. 7.2, the substitution of the incident field for the internal field ultimately results in the phase of the phasors being controlled by the exponential $\exp(i\mathbf{q}\cdot\mathbf{r}_i)$ in Eq. (7.13). Here, however, the phase of the phasors is determined by the product $E_x^{int}\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}_i)$ in Eq. (7.35). Because of the strong coupling throughout the sphere, the magnitude of the phasors will not generally be equal. This means that the phasors are more spread out in the complex plane as compared to the thin ring-like distribution seen in Fig. (7.6).

The objective now is to understand how the phasors given by Eq. (7.35) add together in Eq. (7.34) to cause the features seen in the scattering curves, specifically the power-law crossovers. As in Sec. 7.2, this can be done by studying the way that the distribution of the phasors evolve in the complex plane as q varies. The problem, however, is that now the internal field varies in *magnitude* and phase throughout the sphere, e.g. see Fig. (7.12). This variation in the field results in a complicated phasor distribution that would not be amenable to simple interpretation if presented in the same form as in Fig. (7.6). The problem is resolved by color coding the phasors in the complex plane to match their location inside of the sphere. To explain how this is done, consider Fig. (7.13), which shows the unit disk in the complex plane with the points in the disk color coded. The points are assigned a color and a brightness based upon their phase angle and magnitude, respectively. All points located on the same radial have the same color, whereas the brightness of the color is proportional to its distance from the origin, i.e. to the phasor's magnitude.

Now consider applying this color scheme to the phasors of the same sphere in Fig. (7.12) when the observation point is in the forward direction, i.e. $\hat{\mathbf{r}} = \hat{\mathbf{n}}^{inc}$, which is also $q = \theta = 0$.

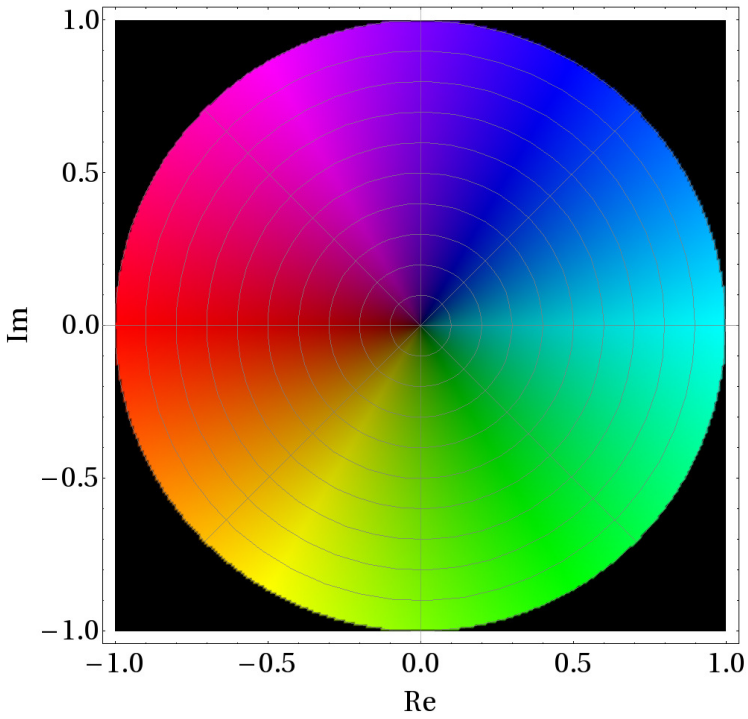

Figure 7.13: Color and brightness assignments for phasors.

Each point in the sphere corresponds to a phasor, and hence the color coding of Fig. (7.13) can be applied to both a plot of the phasors in the complex plane and to the points inside the sphere to which the phasors belong. Figure (7.14) shows an example of this color-coded connection. In plot (a) the same scattering plane cross section through the sphere is shown as in Fig. (7.12), except here the sphere interior is assigned color and brightness as determined by the forward angle ($\hat{\mathbf{r}} = \hat{\mathbf{n}}^{inc}$) phase and magnitude of the factor $E_x^{int} \exp(-ik\hat{\mathbf{n}}^{inc} \cdot \mathbf{r}_i)$. Remember from Eq. (7.35) that this factor determines the location of a wavelet's phasor in the complex plane. Shown in plot (b) is the sphere's phasor distribution in the complex plane with exactly the same color coding. One can compare directly between these plots to see the location inside the sphere that a given phasor in the complex plane represents. For example, consider the cluster of phasors in (b) indicated by the dashed circle. These phasors have the largest magnitudes and correspond to the region in the sphere's interior that is circled in plot (a), c.f. Fig. (7.12).

Notice that the distribution of phasors in plot (b) in Fig. (7.14) are qualitatively clus-

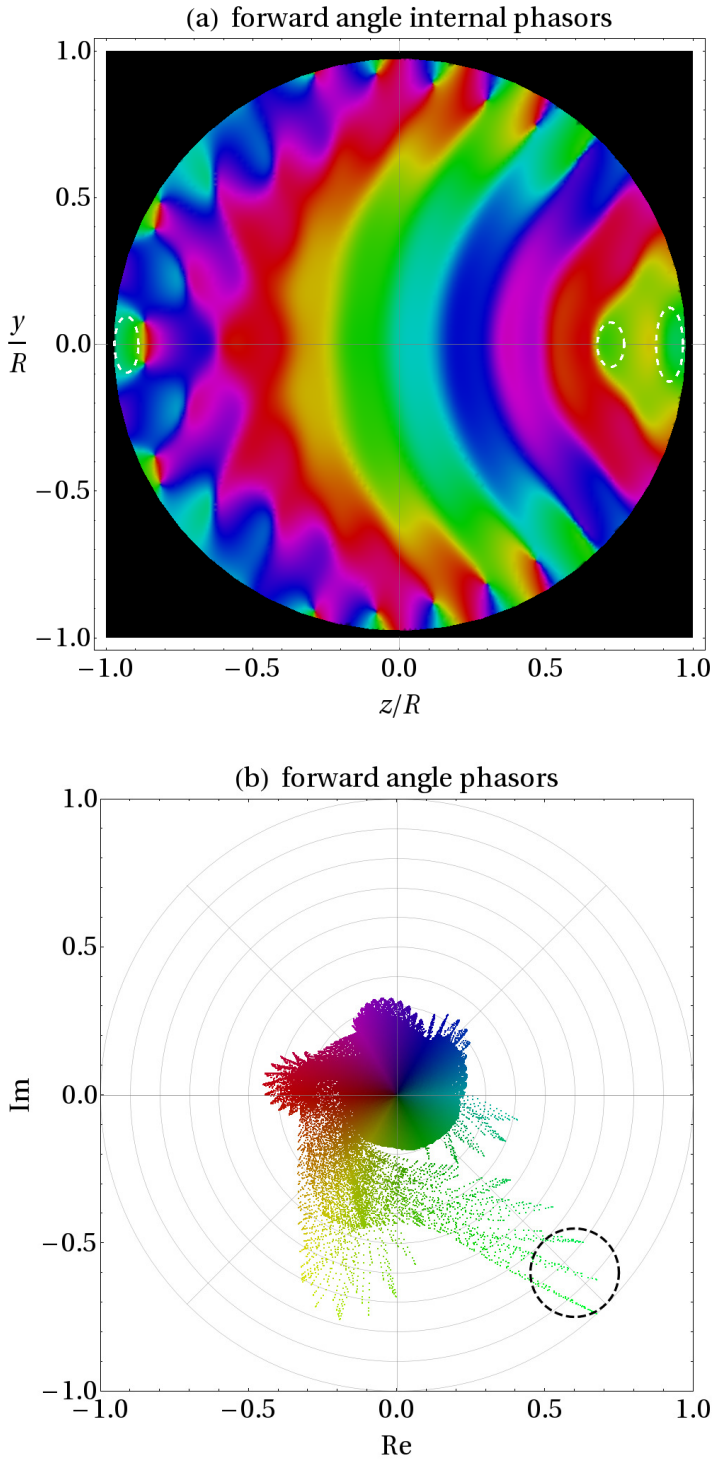

Figure 7.14: Color-coded connection between the interior of the spherical particle (a) shown in Fig. (7.12) and the corresponding phasor distribution (b). The color code is given in Fig. (7.13). The phasors circled in (b) correspond to regions in the sphere with the largest magnitude internal field, which are identified by the circles in (a). Also recall Fig. (7.12).

tered into groups within which the phasors have similar magnitude and phase. An exact quantitative definition for the maximum variation in phase or magnitude defining a group is not necessary. The concept is only intended to enable a qualitative analysis of the distribution's evolution. Now comes an obvious but important realization; the color of the phasors in a given group is roughly the same. This means that as the phasor distribution evolves as q increases from zero, one will be able to track a given phasor group and see how it deforms and moves in the complex plane simply by following its color.

Figure (7.15) shows the normalized q-space scattered intensity curve for the sphere in Figs. (7.12) and (7.14), where $kR = 12$, $m = 1.50 + 0i$ and $\rho = 12$. This scattering curve displays all three power-law regions given in Eq. (7.4). The curve shows a Guinier crossover at $qR \sim 1$ and an extra crossover at $qR \sim \rho$ called the *ρ-crossover*. The first six peaks of the curve are labeled $a - f$ and the phasor distributions for each of these peaks is displayed in Fig. (7.16). The phasor plot a, corresponding to the forward direction $q = 0$, is exactly the same plot as seen in Fig. (7.14).

Now compare phasor plot b to a. Between these plots, the scattering angle advances from zero to the Guinier crossover, and one can see that as a consequence the phasors have slightly rotated resulting in the twisted appearance of the plot. Remember that the colors are assigned to the phasors when the scattering angle is zero. The phasors are not recolored as they rotate to other positions in the complex plane. Recall from Sec. 7.2 that spreading of the phasors in the complex plane causes cancellation between them and consequently reduces the scattered intensity. Now return to the phasor-plot a in Fig. (7.16). Since the phasors are spread around the origin, there is a large degree of cancellation among them, even though the scattering angle is zero. This is contrary to the $\rho < 1$ case in Sec. 7.2. The cancellation of phasors in plot a occurs between different phasor groups, i.e. between different colors. This is *inter-group* phasor cancellation. As the scattering angle increases from zero to the Guinier crossover, the phasors spread, which enhances inter-group cancellation leading to a decrease in the scattered intensity. The magnitude of the enhancement in inter-group

cancellation becomes significant only when the scattering angle is large enough to advance the phase of the exponential $\exp(-ik\hat{\mathbf{r}} \cdot \mathbf{r}_i)$ in Eq. (7.35) relative to kR. This shows that the Giunier crossover occurs in the $\rho > 1$ case for the same reason that it does in the $\rho < 1$ case; wavelets located at the "caps" of the sphere in the $\pm\hat{\mathbf{q}}$ direction begin to cancel out via destructive interference.

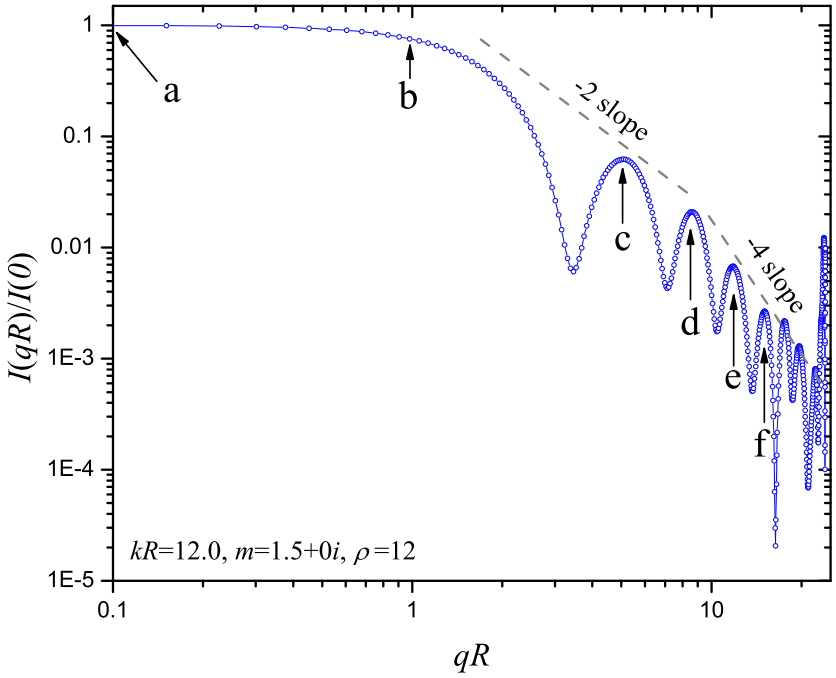

Figure 7.15: Normalized scattered intensity curve for the sphere in Figs. (7.12), (7.14), and (7.16). The points labeled $a - f$ correspond to the phasor plots, labeled likewise, in Fig. (7.16).

Next, by comparing the qualitative structure of the phasor plots *b*-*d* one can see that the phasors continue to spread in phase angle as the scattering angle increases. Very careful inspection of these distributions reveals that by the time the ρ-crossover is reached, the phasors within a given group have spread by an amount of nearly π in phase. This can be seen by the smearing of a given color over an angular range of π in plot *d*. Once the phasors of a given group spread to this extent, they begin to cancel each other, and hence constitute *intra-group* phasor cancellation. Notice though, that inter-group cancellation is still active.

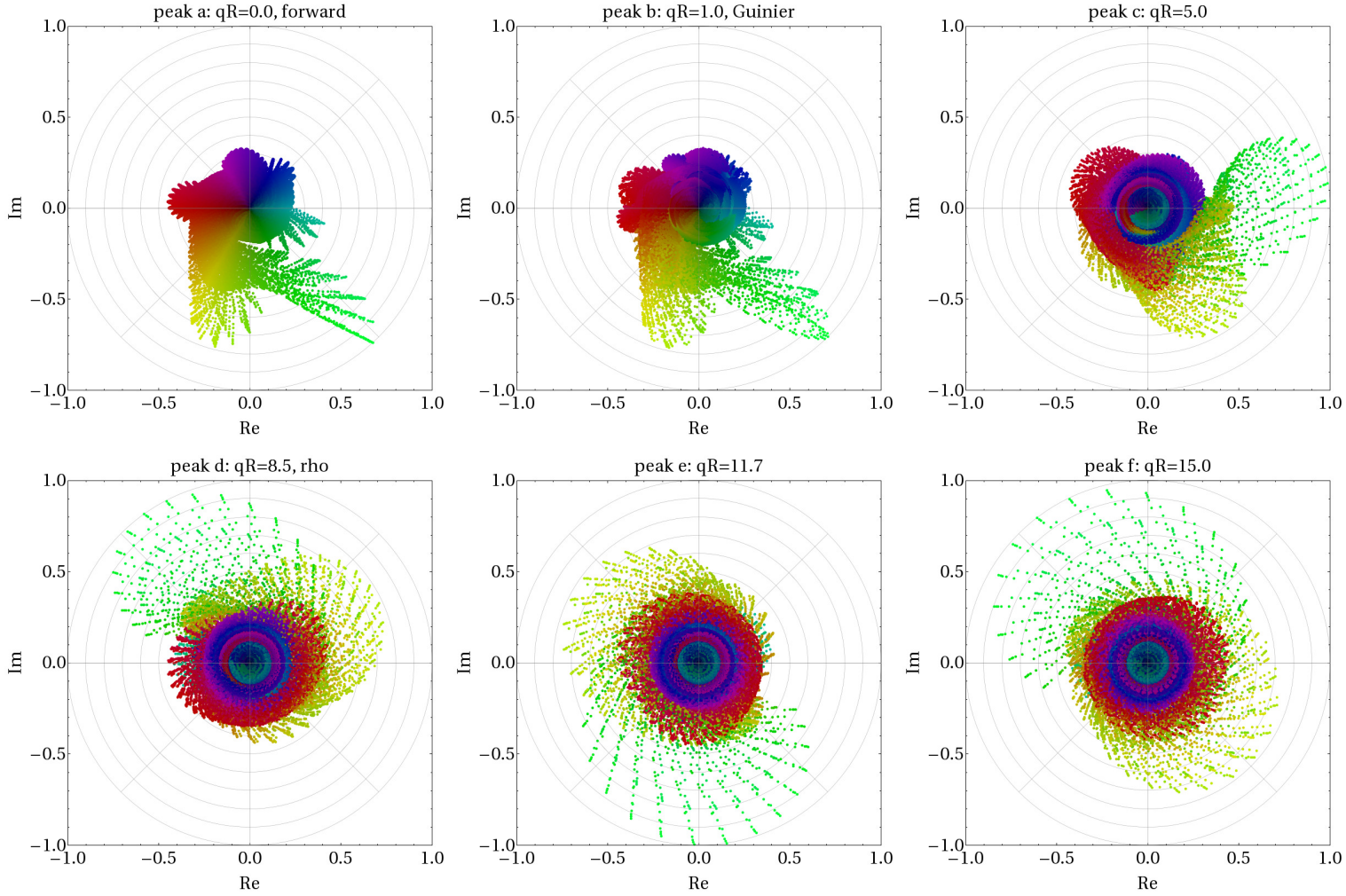

Figure 7.16: Color-coded phasor distributions for the points on the scattering curve in Fig. (7.15) identified by $a - f$. Distribution a is exactly the same in plot (b) in Fig. (7.14), which is in turn related to plot (a) in that figure. The phasor color codes are assigned according to the distribution at $\theta = 0$, i.e. plot a, c.f. Fig. (7.13).

As mentioned above, the onset of inter-group cancellation corresponds to the beginning of destructive interference between wavelets separated by the width of the sphere along $\pm\hat{\mathbf{q}}$. This is why the Guinier crossover is related to the sphere radius. Then, it is reasonable to expect that the onset of intra-group cancellation at the ρ-crossover signifies a destructive interference effect associated with some other physical aspect of the particle.

Before the ρ-crossover, the contributions to the scattered intensity comes from the phasors of different groups that cancel each other poorly. To explain this in more detail, return to phasor plot *a* in Fig. (7.16) and find the phasor group, colored green, that consists of the largest magnitude phasors. This group is poorly canceled by other groups due to its greater than average magnitude, see plots *a*-*d*. It is only after the ρ crossover is reached, and intra-group cancellation begins, that the contribution to the scattered intensity from this group is substantially decreased. One can now return to Fig. (7.14) and see that this (green) group belongs to the region in the sphere where the internal field magnitude is greatest. Remember that the brightness of the colors indicates the phasor magnitude. Regions like this of localized enhanced field magnitude are often referred to as hot spots. The cause of the hot spots is ultimately due to the coupling between the particle's wavelets, which one can understand by realizing that there are no hot spots for the $\rho < 1$ case where refraction, or coupling, is weakest. Given this, one can conclude that *the significance of the ρ crossover is that it identifies the scattering angle at which the phase of the wavelets has advanced enough that destructive interference begins between the wavelets within the particle's hot spots.* This shows why the ρ crossover depends on ρ since ρ is a measure of the total degree refraction in the sphere, which is partly the cause for the hot spots.

From the qualitative power-law behavior of the scattering curves in Eq. (7.4), one can *empirically* identify "length scales" that are "sensed" by the scattering wave vector q:

$$\begin{aligned} \frac{1}{q_G} &\simeq R, \quad \text{Guinier crossover} \\ \frac{1}{q_\rho} &\simeq \frac{R}{\rho}, \quad \rho\text{-crossover} \end{aligned} \tag{7.36}$$

where q_G and q_ρ denote the location of the crossovers in the q-space scattering curves, eg. recall Fig. (7.3). The discussions above explain that the Guinier crossover is due to destructive interference of the wavelets at opposite ends of the sphere. Then, the length scale associated with this crossover must clearly depend on sphere size R. Notice from Eq. (7.36) that the length scale associated with the ρ crossover is independent of sphere size since R cancels in q_ρ via Eq. (7.3). The length associated with the ρ crossover relates to the features in the internal field due to refraction, i.e., hot spots, and hence depends on the wavelength and refractive index. Notice that this is not inconsistent with the ρ universality seen in Fig. (7.3) because the curves there are scaled by their respective values of R, i.e., they are plotted as a function of qR not q.

As a final topic, recall that a major difference between the phasor distribution at zero scattering angle in the $\rho < 1$ and $\rho > 1$ cases is that the former is clustered tightly around the real axis while the latter is spread completely around the origin of the complex plane. This can be seen by comparing phasor plots a in Figs. (7.6) and (7.16). As mentioned above, this spread around the complex plane for the $\rho > 1$ sphere shows that a substantial degree of phasor cancellation occurs even at zero scattering angle. In contrast, the scattering angle has to advance past the Guinier crossover before cancellation begins for the $\rho < 1$ sphere. This means that a sphere with $\rho > 1$ has less scattered intensity in the forward direction as compared the same sphere treated in the RDG approximation, see [79]. The above analysis demonstrates that the cause of this is refraction. In the forward direction the phase of phasors is determined by $E_x^{int} \exp(-ik\hat{\mathbf{n}}^{inc} \cdot \mathbf{r}_i)$ in Eq. (7.35). The exponential has the same functional form as the incident wave, cf. Eq. (3.1). Then the phase of this field-exponential product represents the difference in phase between the internal and incident wave. If ρ is large, refraction is strong, and the relative phase varies extensively throughout the sphere interior. This spreads the phasors around the origin in the complex plane as seen in a of Fig. (7.16) and causes the RDG-relative reduction in forward scattered intensity seen in [79].

Chapter 8

Fractal-Like Aggregates

"I am a dangerous creature."
-A. A. Lushnikov, AAAR 2006, St. Paul, MN.

Fractal aggregates are irregularly shaped particle composites that form under special physical conditions. They can occur, for example, as a result of incomplete combustion of hydrocarbon fuels [81]. Consequently, there is a substantial amount of work in the literature devoted to understanding the scattering behavior of such carbon aggregates, especially as it relates to the radiative forcing effects that such aggregates may have in the atmosphere [74, 82]. The literature on this topic is so extensive that it cannot be reasonably be reviewed here; the interested reader is referred to [74]. The reason that fractal aggregates are of interest here is because they offer an opportunity to apply the microphysical model to a complicated particle shape. In the process of doing so, one will see the major qualitative features of the fields that reside inside these complex-shaped aggregates and how this relates to their scattering properties. The following will also be able to weigh in on the long-standing debate regarding the validity of the RDG approximation for fractal aggregates.

8.1 Morphology

Image (a) in Fig. (8.1) shows a transmission electron microscope picture of part of a soot aggregate captured in an open-burning acetylene flame [83]. This aggregate is composed of

an assembly of smaller, roughly spherical, monomers. The monomers are carbon particles of roughly the same size. Although the details of the aggregate's structure depends on complicated inter-monomer interactions, nearly all aggregates like this one are known to obey a statistical scaling law, Eq. (8.1) below [74]. Models of aggregate formation have been developed that also follow this scaling law [84]. Image (b) in Fig. (8.1) shows a simulated aggregate generated by one such model; diffusion limited cluster aggregation (DLCA), see [84, Sec. 2.10].

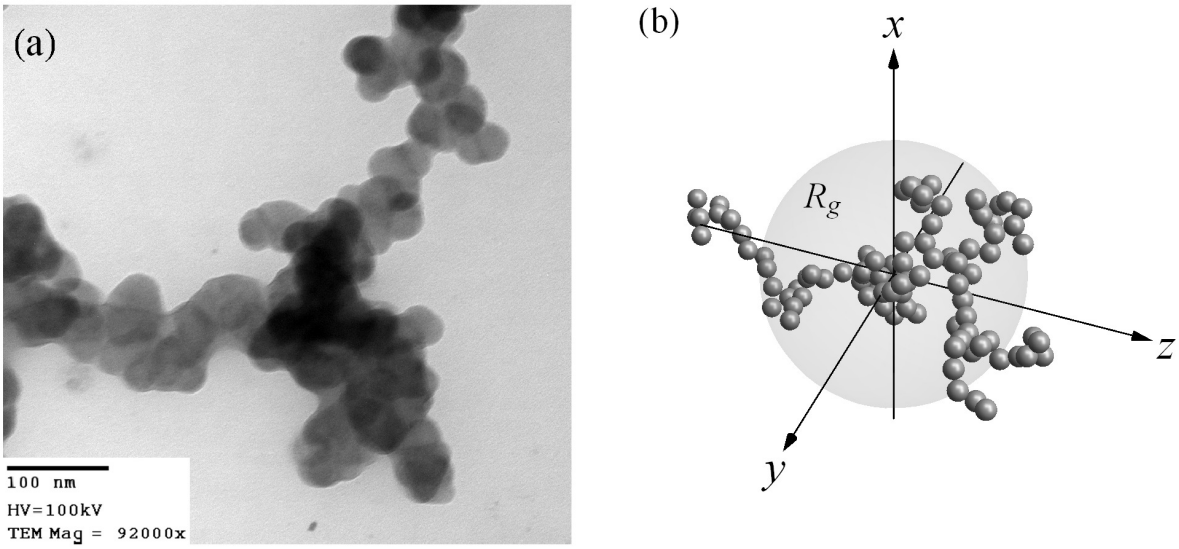

Figure 8.1: Open-flame acetylene soot aggregate (a), and DLCA simulated aggregate (b) with $N_m = 99$ monomers and a fractal dimension $D_f \sim 1.8$.

The DLCA method produces aggregates that obey a scaling law involving quantities that describe the gross aggregate morphology,

$$N_m = k_o \left(\frac{R_g}{R_m} \right)^{D_f}, \tag{8.1}$$

where N_m is the total number of monomers in the aggregate, k_o is a proportionality constant, R_g is the aggregate radius of gyration, R_m is the monomer radius, and D_f is the fractal dimension. Aggregates that obey this scaling law possess a fractal-like character in that they show self-similar features in their overall morphology over a range of length scales. This is different than a pure fractal, which is self similar in morphology on all length scales

[85, p. 11]. The parameters in Eq. (8.1) that influence the shape of the aggregate are the prefactor k_o and the fractal dimension D_f. Reference [82] shows images of aggregates for different values of these parameters. For the DLCA aggregates considered below, the fractal prefactor is a constant with a value of $k_o \simeq 1.3$ and $D_f \simeq 1.8$. Variations in the fractal dimension D_f correspond to major qualitative changes in the aggregate morphology; $1 \lesssim D_f \lesssim 2$ gives "stringy" linear-like aggregates that become more compact as D_f increases up to a maximum of three. Aggregates that form due to incomplete combustion like the one in Fig. (8.1) usually correspond to $D_f \simeq 1.8$ [74]. The parameters R_g, R_m and N_m relate to the aggregate size. The radius of gyration R_g provides a measure of the overall size of the aggregate

$$R_g = \frac{1}{\sqrt{N_m}} \sqrt{\sum_{i=1}^{n_m} |\mathbf{r}_i - \mathbf{r}_{com}|^2}, \tag{8.2}$$

where $\mathbf{r}_i$ is the location of a monomer and $\mathbf{r}_{com}$ is the location of the aggregate's center of mass [80, p. 24-25].

8.2 Aggregates in Q-space

If an aggregate's fractal dimension is less than three, like $D_f \simeq 1.8$ for open-flame soot, the monomers are arranged with roughly two nearest neighbors on average, recall Fig. (8.1). Because the inter-monomer spacing is greater than the spacing between the equal-sized volume elements of a compact particle of the same R_g, one might expect that the coupling between monomers is negligible and only the influence of the incident wave is important. If this is so, then the RDG approximation may be valid. The work following this section will quantitatively study the validity of this approximation.

If the RDG approximation is valid and the observation point is confined to the half contour $\mathcal{C}_h$ shown in Fig. (7.1), then the considerations in Sec. 7.2 show that the normalized q-space scattered intensity is given by

$$\frac{I(\mathbf{q})}{I(0)} = \left| \sum_i \exp(i\mathbf{q}\cdot\mathbf{r}_i) \right|^2 \quad \mathbf{r} \in \mathcal{C}_h, \tag{8.3}$$

where Eqs. (7.13) and (7.15) have been used.

One typically assumes that there are many similar aggregates in arbitrary random orientations in the scattering volume during the course of a measurement. Then, the scattering quantities for such a system of aggregates can be modeled as the ensemble average of the scattering quantities associated with a *single* aggregate in all orientations, recall Sec. 5.4. As a result, the scattered intensity of Eq. (8.3) is dependent on q only, rather that $\mathbf{q}$. With this simplification, [74] shows that the q-space scattered intensity has the power-law form

$$I(q) \propto \begin{cases} N_m^2 & q \lesssim 1/R_g \\ N_m^2 (qR_g)^{-D_f} & 1/R_g \lesssim q \lesssim 1/R_m \\ N_m^2 (qR_g)^{-4} & q \gtrsim 1/R_m. \end{cases} \tag{8.4}$$

Equation (8.4) demonstrates that the scattered intensity for an aggregate has a forward angle maximum that is proportional to the number of monomers in the aggregate. Then, following the Guinier crossover corresponding to the radius of gyration R_g, the scattering curve enters a power-law region where the curve decreases with increasing q with a power-law exponent roughly equal to $-D_f$. If q increases enough, a second crossover occurs corresponding to the monomer radius. Q-space scattered intensity plots are presented in [74, 84, 86] and show good qualitative agreement between the behavior described in Eq. (8.4) and measurements performed on an aggregating aerosol systems.

Because of the highly complex shape of aggregates, the simplicity of the RDG approximation makes it attractive for calculation of their scattering quantities. Only recently has this problem been solved exactly, in a numerical sense, using the superposition T-Matrix method [82]. Unfortunately, since this T-Matrix treatment is designed to calculate the aggregate's far-field scattering quantities, it does not readily provide direct insight to behavior of the monomer internal-fields. Consequently, the validity of the RDG approximation is not directly corroborated (or not) by this T-Matrix work.

8.3 Aggregates in the DDA

The derivation of the q-space power-law scattered intensity described in Eq. (8.4) fundamentally relies on the RDG approximation along with orientational averaging. In this regard, it is important to establish how well the RDG approximation relates to the actual monomer-internal fields[1]. The following will use the DDA to calculate the far-field Stokes parameters of a DLCA simulated fractal aggregate. In the process, the DDA will provide the internal field, enabling a direct comparison with the RDG approximation and hence quantify the validity of the approximation. The reader should note that the following is not intended to constitute an extensive study of scattering from aggregates; rather, the emphasis is on the monomer internal-fields and the effects attributed to orientational averaging.

8.3.1 DLCA Aggregates

The simulated aggregates considered in much of the following is shown in (b) of Fig. (8.1). Since this aggregate is created using DLCA, it has a fractal prefactor of $k_o \simeq 1.3$ and fractal dimension of $D_f \simeq 1.8$ [84]. Shown in the figure along with the aggregate is an opaque sphere indicating the size of the aggregate's radius of gyration R_g. The monomer size parameter is $kR_m = 0.41$ in most of the following, and the refractive index[2] is $m = 1.57 + 0.6i$. Both of these values are representative of carbon soot aggregates illuminated by visible light [74]. For example if $\lambda = 532$ nm, $R_m = 34.5$ nm. One can compare this to the aggregate in (a) of Fig. (8.1) and see rough agreement with the actual soot monomer sizes.

The DDA is applied by superimposing the aggregate on the DDA lattice and setting to zero any dipoles that reside outside of the monomers. The fineness of the lattice determines how well the aggregate shape is represented, i.e., the shape errors of Sec. (5.5), and can be quantified by the number of dipoles per monomer. The most computationally demanding aspect of the DDA calculations is orientational averaging, which competes against

[1] One should note that work is hardly the first to address this issue, see [74].

[2] There is some debate over what the refractive index is for carbon soot in the optical frequencies, see [74].

the computational demands of increasing the lattice fineness. From Sec. 5.4, the orientational averaging procedure requires that the entire DDA matrix equation (5.15) be solved for each aggregate orientation. These orientations are described by the Euler angles using the rotation matrices of Eqs. (5.25), (5.27), and (5.30).

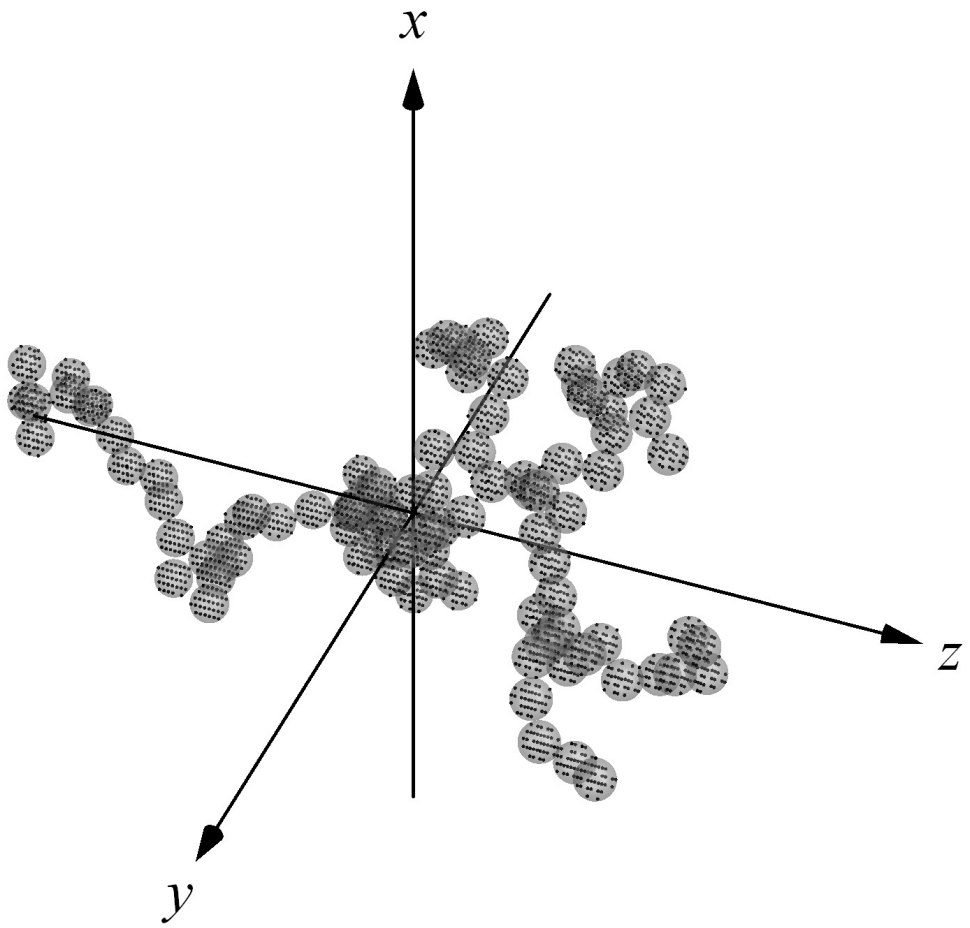

Figure 8.2: Same DLCA-generated fractal aggregate as in (a) of Fig. (8.1), except here the DDA lattice sites are shown inside of the monomers. In this particular case, the total number of lattice sites is $N_{dda} = 3323$ corresponding to about 36 sites per monomer.

The primary scattering quantities of interest are the Stokes parameters of Eqs. (1.47)-(1.50), and are found from the DDA far-field scattering amplitude of Eq. (5.23). For each aggregate orientation, these parameters are calculated for points in the LCS on the surface of the large imaginary sphere $\mathcal{S}_{en}$ of Sec. (6). The orientational average of a Stokes parameter

is computed via the discretized approximate ensemble average of Eq. (5.32). Whether for a single orientation, or for the ensemble average, the far-field scattered intensity curves (i.e. the Stokes parameter I), are calculated only for observation points on the half-contour $\mathcal{C}_h$ shown in Fig. (7.1). This enables comparison to the q-space power-law behavior of Eq. (8.4).

Several analysis techniques are used to study the monomer internal field and compare it to the RDG expectation. The relative RDG magnitude deviation S_i^{RDG} is defined as

$$S_i^{RDG} = \frac{1}{N_{ori}} \sum_{n=1}^{N_{ori}} \frac{|\mathbf{E}_n^{int}(\mathbf{r}_i)| - |\mathbf{E}_n^{inc}(\mathbf{r}_i)|}{|\mathbf{E}_n^{inc}(\mathbf{r}_i)|}. \tag{8.5}$$

Here $\mathbf{r}_i$ is the location of the i^{th} DDA lattice site inside of the aggregate and the index n denotes the n^{th} member of a total of N_{ori} aggregate orientations. This quantity gives the average deviation of the internal electric field at a lattice site from the RDG expectation. If $S_i^{RDG} = 0$, the field at that site is equivalent to what the RDG approximation predicts. Values greater than zero indicate a deviation in magnitude of the average field at that lattice site relative to the corresponding magnitude of incident wave. Visualizations of S_i^{RDG} will be given either as a histogram showing the number of lattice sites with deviations in certain ranges, or as color-coded renderings of the actual aggregate.

To analyze the directional character of the monomer internal field, each lattice site in the aggregate is assigned a unit vector $\hat{\mathbf{n}}_i^{int}$ representing the direction of the internal field there. Each aggregate orientation results in a different unit vector for the site, and is then averaged as

$$\mathbf{n}_i^{int} = \sum_{n=1}^{N_{ori}} \frac{\text{Re}\left\{\mathbf{E}_n^{int}(\mathbf{r}_i)\right\}}{\left|\text{Re}\left\{\mathbf{E}_n^{int}(\mathbf{r}_i)\right\}\right|}. \tag{8.6}$$

To visualize these average field directions a point is plotted on the surface of the unit sphere in the LCS, where each point corresponds to the unit vector $\hat{\mathbf{n}}_i^{int}$ given by Eq. (8.6). Since the incident electric field is polarized along the x-axis, these points will be clustered around the x-axis poles of the unit sphere if the RDG approximation holds well for an given orientational ensemble. A dispersion of points away from the poles indicates a directional

deviation from the RDG approximation. One should note that the magnitude and direction deviation quantities given in Eqs. (8.5) and (8.6) are not intended to represent observables. These quantities are directly proportional to the field as opposed to the (observable) Stokes parameters.

Figure (8.2) shows the aggregate used in the DDA calculations resulting in Figs. (8.3)-(8.9) below. Here the monomers are shown as translucent spheres. The dipole lattice sites inside of the monomers are displayed as black dots. The total number of monomers in this aggregate is $N_m = 99$ and the total number of lattice sites is $N_{dda} = 3323$. This gives an average of 36 dipole sites per monomer; the slight variation in this number from monomer to monomer is due to the lattice shape error. The monomer-relative size of the aggregate is $R_g/R_m = 12.9$, corresponding to a DDA size parameter of $kR_g = 5.4$, where λ equals 30 DDA lattice spacing units d.

Shown in Fig. (8.3) is the aggregate with each monomer assigned a color corresponding to the relative RDG deviation integrated over the volume of the monomer *for a single aggregate orientation*; this is not an orientationally averaged field deviation. Using the argument of the sum in Eq. (8.5), this integrated deviation is given by

$$\frac{1}{V_m}\int_{V_{m,i}} \frac{|\mathbf{E}^{int}(\mathbf{r}_i)| - |\mathbf{E}^{inc}(\mathbf{r}_i)|}{|\mathbf{E}^{inc}(\mathbf{r}_i)|}\,\mathrm{d}\mathbf{r}, \tag{8.7}$$

where V_m is the volume of a monomer[3] and the notation $V_{m,i}$ denotes integration over the i^{th} monomer of the aggregate. If this quantity is zero then the field throughout the monomer is the same as the RDG expectation whereas if it is greater (less) than zero, the internal field is greater (less) than RDG. One can see from this figure that some monomers have more than $\pm 25\%$ integrated deviation from the RDG expectation, and most monomers are at least $\pm 10\%$ deviated from RDG.

Figure (8.4) shows histogram plots of the number of dipole lattice sites with S_i^{RDG} values within a range of deviation as shown, recall Eq. (8.5). Each plot shows the RDG magnitude deviation for different numbers of orientations in the average. Plot (a) shows the

[3]Remember that all the monomers are the same size.

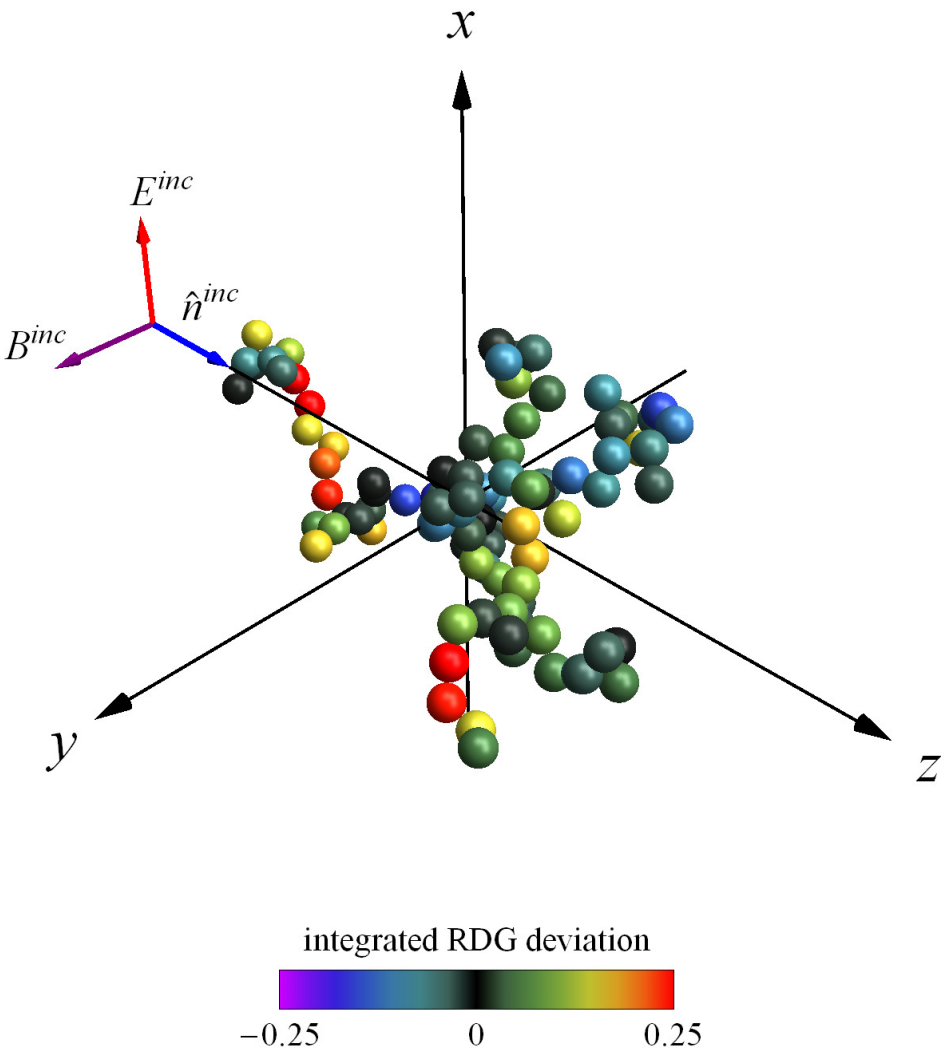

Figure 8.3: Same aggregate as in Fig. (8.2) with each monomer color coded according to Eq. (8.7).

deviation for a single aggregate orientation $N_{ori} = 1$ and corresponds to the same aggregate orientation in Fig. (8.3). Plots (b) and (c) show the deviation distribution for $N_{ori} = 10$ and $N_{ori} = 500$, respectively. Comparison between these plots shows that increasing the number of orientations in the average, Eq. (8.5), results in a narrowing of the distribution. The overall width of the deviation for $N_{ori} = 50$ is roughly $\pm 5\%$. Moreover, the shape of the distributions shows that the internal field magnitude tends to be slightly greater than the RDG expectation as opposed to less than it. Each distribution is peaked near a deviation of zero indicating that the largest fraction of lattice sites have an internal field magnitude the agrees well with RDG, within $\pm 1 - 3\%$.

Now consider the RDG internal-field direction deviation of Eq. (8.6). Figure (8.5) shows the unit sphere in the LCS with the directions $\hat{\mathbf{n}}_i^{int}$ of Eq. (8.6) represented as points on the sphere. The three plots $(a) - (c)$ show the directions for $N_{ori} = 1$, $N_{ori} = 10$, and $N_{ori} = 500$ orientations in the average. The single orientation plot corresponds to Figs. (8.3) and (8.4). One can see in (a) that most of the lattice sites have internal field directions that are close to the $\pm\hat{\mathbf{x}}$ directions, which is consistent with the RDG expectation. Interestingly, as more orientations are included in the average in Eq. (8.6), the dispersion of points appear essentially the same as for the single orientation.

In Fig. (8.6) are plots like Fig. (6.8) showing the vibration ellipses and corresponding ellipticity of the far-field scattered wave on the surface of the large imaginary sphere $\mathcal{S}_{en}$, recall Secs. 1.6 and 6.1.3. The ellipticity is calculated from the Stokes parameters via Eqs. (1.51) and (1.52) and is indicated on $\mathcal{S}_{en}$ by the gray shades. The handedness, or rotation, of the polarization is indicated by color; red denotes left-handed and blue denotes right-handed rotation. As with Figs. (8.4) and (8.5), three plots are shown $(a) - (c)$ corresponding to $N_{ori} = 1$, $N_{ori} = 10$, and $N_{ori} = 500$ orientations in the *ensemble* average of the Stokes parameters. Plot (a) reveals that the polarization state of the scattered wave is mostly elliptical, including at points on the $\mathcal{C}_h$ contour[4]. An especially important direction is

[4]Recall that this contour is defined by the intersection of the imaginary spherical surface $\mathcal{S}_{en}$ with the y-z plane.

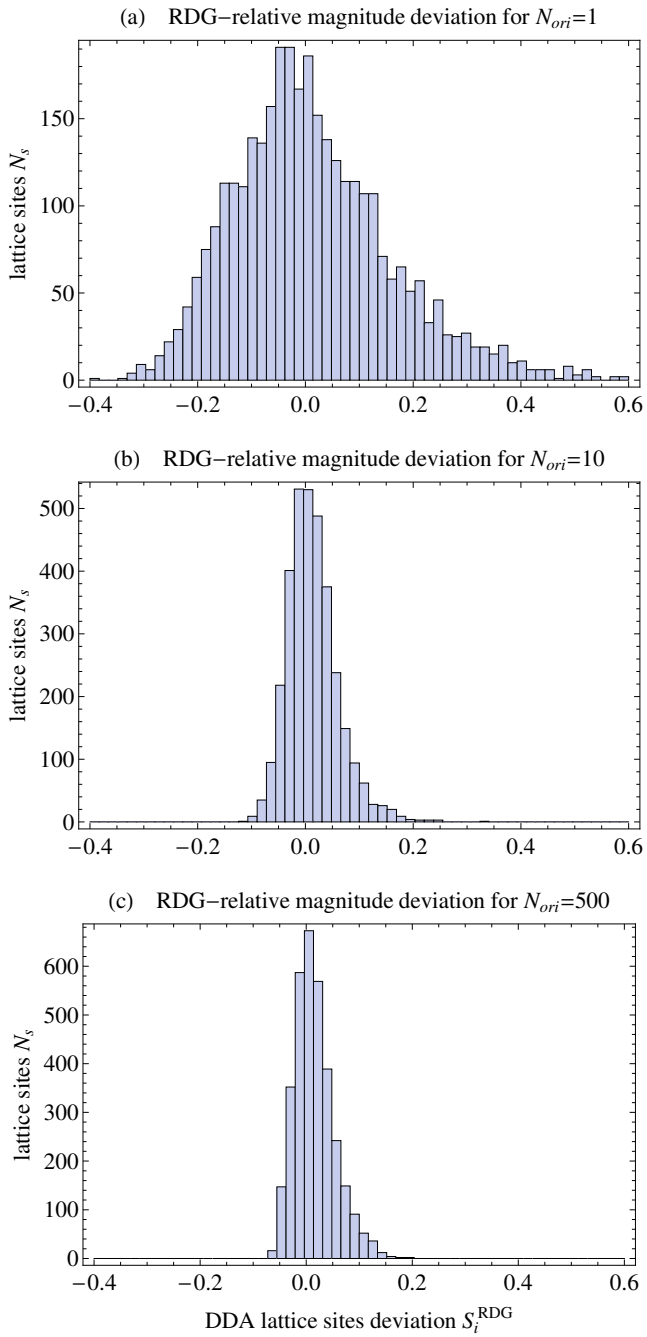

Figure 8.4: Histogram plots of the RDG-relative internal field-magnitude deviation given by Eq. (8.5). Plot (a) corresponds to only one orientation, the same as shown in Fig. (8.3), while plots (b) and (c) correspond to $N_{ori} = 10$ and $N_{ori} = 500$ orientations, respectively.

the backscattering direction where one can see that the scattered wave is elliptical. This is consistent with the nonspherical shape of the aggregate in the context of the phasor cancellation explanation of the polarization state given in Secs. 6.1.4 and 6.2, since the aggregate has no reflection symmetry about both the Π_x and Π_y planes. Plots (b) and (c) demonstrate that including more orientations in the ensemble average of the Stokes parameters results in the polarization state becoming more linear overall.

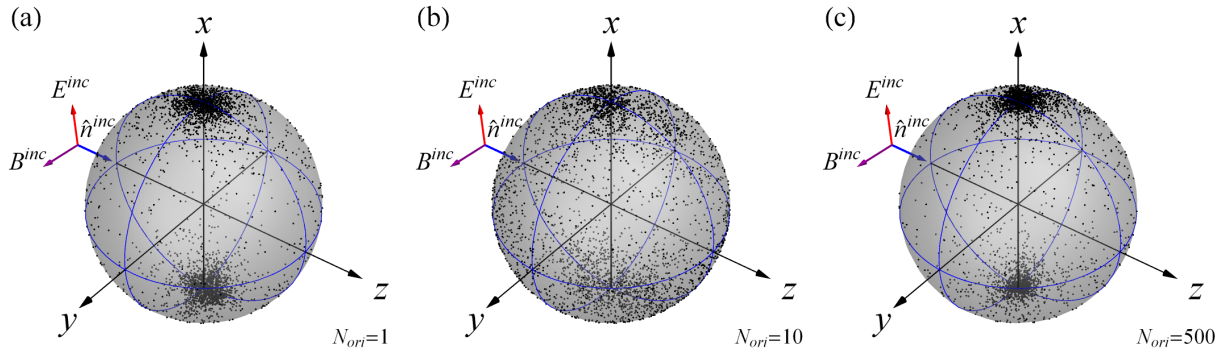

Figure 8.5: Plots of the RDG-relative internal field-direction deviation of Eq. (8.6) for $N_{ori} = 1$, $N_{ori} = 10$ and $N_{ori} = 500$ orientations, plots (a), (b), and (c), respectively. The aggregate is the same as in Figs. (8.2)-(8.4). The orientation of the aggregate for (a) is the same as shown in Fig. (8.3).

The discussion in Sec. 6.1.2 regarding Eq. (6.36) shows that if the internal field is uniform in direction, the scattered wave must be linearly polarized in all directions. Consequently, if the RDG approximation is valid for an aggregate, the polarization state should be linear over all directions, which is roughly what is seen in (c) of Fig. (8.6). This appears to show that *the reason that the polarization state is mostly linear is due to orientational averaging*, cf. [86, 87]. Hence, a polarization measurement along the $\mathcal{C}_h$ contour for a *single* aggregate in a *single fixed orientation* should find points where the scattered wave is highly elliptical, as can be seen (a) in Fig. (8.6). However, the same measurement on a swarm of randomly oriented aggregates should show a mostly linear polarization.

Because these DDA calculations do not analytically average the scattering quantities, one must somehow establish if enough orientations are included such that the discrete ensemble average converges to its analytical value. There are multiple ways to do this as mentioned

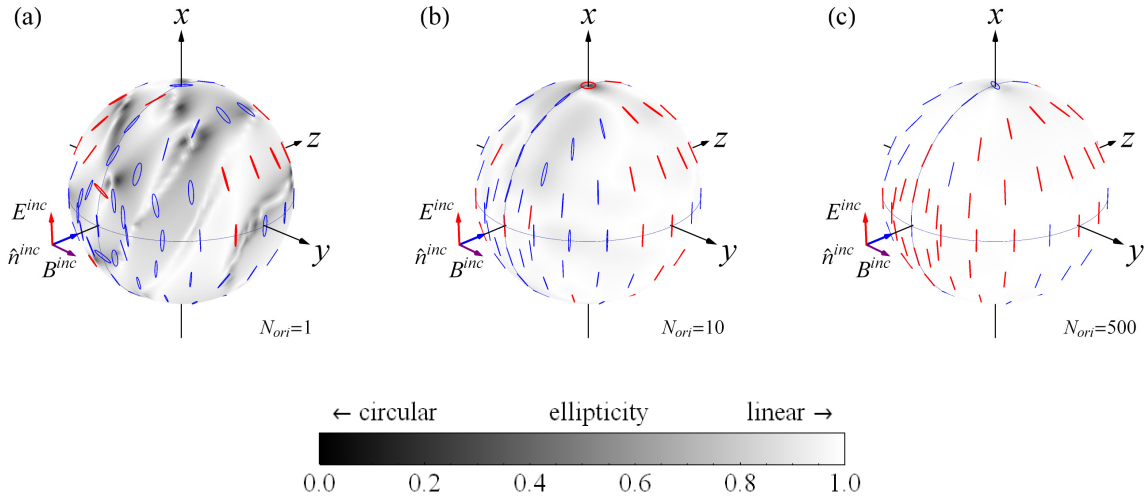

Figure 8.6: Plots like Fig. (6.8) showing the polarization state of the far-field scattered wave. The aggregate is the same as in Fig. (8.2)-(8.5), and plots (a), (b), and (c) show the polarization state after $N_{ori} = 1$, $N_{ori} = 10$ and $N_{ori} = 500$ orientations, respectively. The orientation of the aggregate in the (a) is the same as shown in Fig. (8.3).

in Sec. 5.4. Because one of the interests in this work is to compare to the q-space behavior in Eq. (8.4), the convergence is established from the evolution of the scattering curves. Figure (8.7), plot (a) shows the normalized intensity curves as a function of qR_g along the half-contour $\mathcal{C}_h$. Each of the curves corresponds to the orientational ensemble average of the I Stokes parameter for the number of aggregate orientations indicated. Plot (b) shows the same as (a) except with only the $N_{ori} = 1$ and $N_{ori} = 500$ curves. One can see from these plots that the orientational average converges well by $N_{ori} = 500$ and follow the q-space behavior of Eq. (8.4) rather well.

One of the predictions of the RDG approximation is that a fractal aggregate's forward scattered intensity is proportional to the Rayleigh scattering cross section C_{Ray}^{sca} times the square of the number of monomers, i.e. $I(0) = N_m^2 C_{Ray}^{sca}$, where the Rayleigh cross section is [74]

$$C_{Ray}^{sca} = \frac{8}{3}\pi R_m^2 (kR_m)^4 \left|\frac{m^2-1}{m^2+2}\right|^2 . \tag{8.8}$$

To test this prediction, Fig. (8.8) shows the orientational ensemble averaged scattering curve as in Fig. (8.7) for $N_{ori} = 200$, except normalized by C_{Ray}^{sca}. Three monomer size parameters

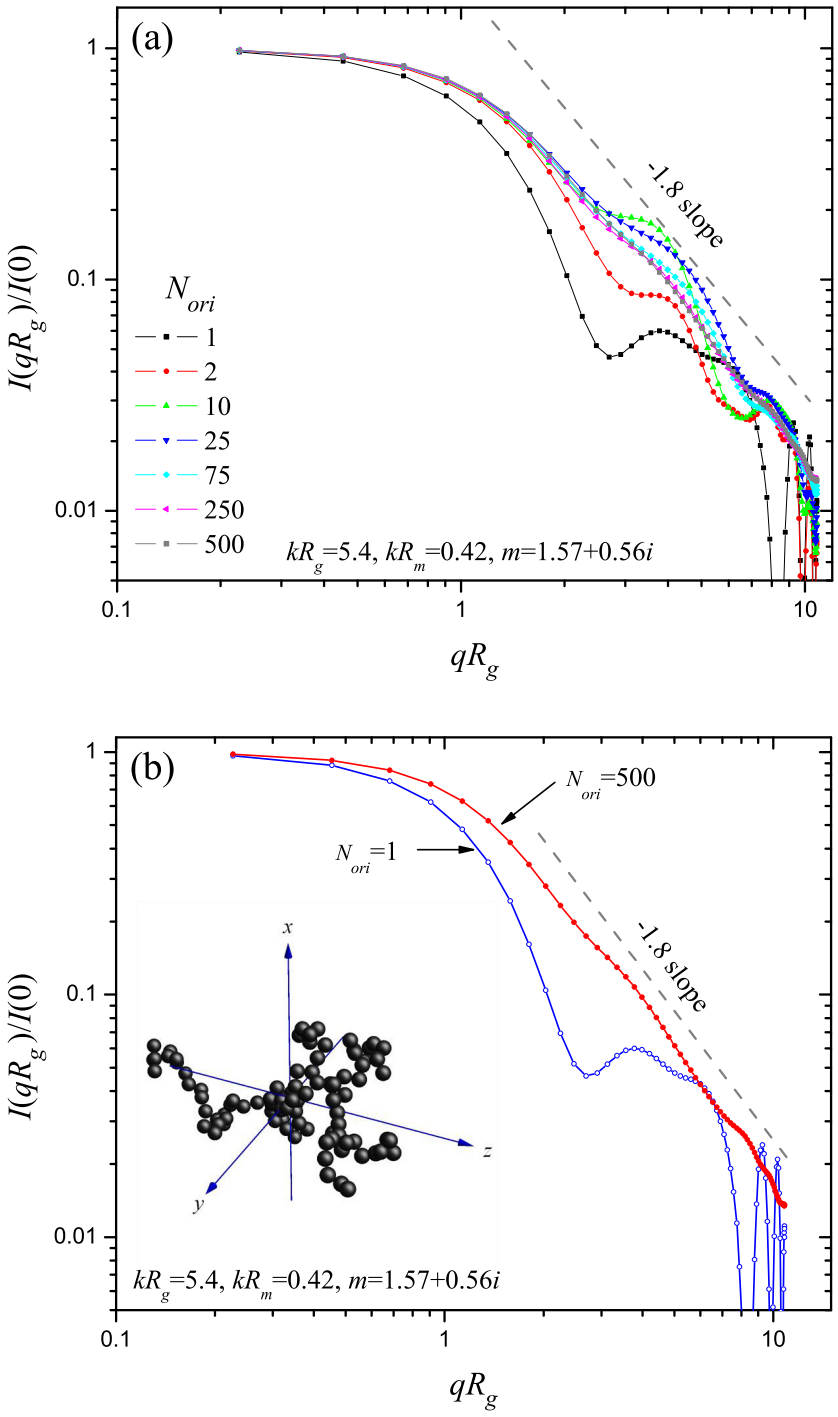

Figure 8.7: Normalized scattered intensity for the aggregate in Figs. (8.2)-(8.6) as a function of qR_g. Plot (a) shows the evolution of the curves as the number of orientations varies from $N_{ori} = 1$ to $N_{ori} = 500$. For clarity, plot (b) shows only the $N_{ori} = 1$ and $N_{ori} = 500$ curves in (a) along with the aggregate.

kR_m are considered. The curve corresponding to $kR_m = 0.21$ is scaled to unity and the other curves are scaled by the same factor; hence, if the expectation is correct, all the curves should lie on one another. One can see that the prediction works well, to within about $10-15\%$, which is roughly consistent with the deviation from the RDG approximation seen in (c) Fig. (8.4).

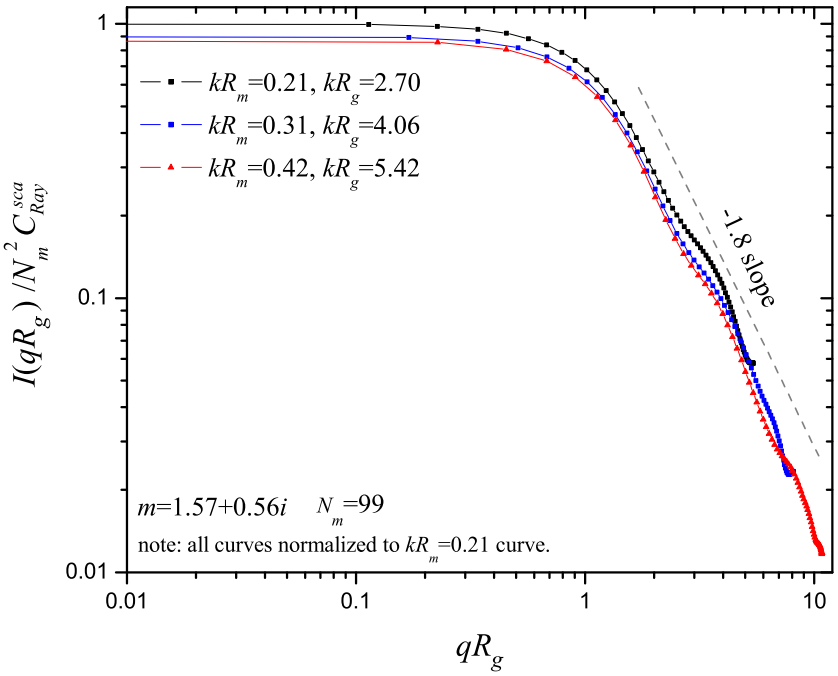

Figure 8.8: Scattered intensity curves for the same aggregate as in Figs. (8.2)-(8.7), except here the monomer size is varied as shown. Each curve is normalized by $N_m^2 C_{Ray}^{sca}$ as calculated from Eq. (8.8) for the specific value of monomer size. Note that the curve for the $kR_m = 0.21$ is scaled to unity and the same scale factor is applied to the other curves.

The aggregates considered so far are too small for the monomer crossover given in Eq. (8.4) to appear. This is because q can reach a maximum value of $2k$ and hence can be insufficiently large to sense the crossover at $q \simeq 1/R_m$. Remember that λ cannot be arbitrarily decreased in the DDA due to the limited fineness of the DDA lattice. Figure (8.9) shows the normalized q-space curves for an aggregate of $kR_g = 13.5$, which is large enough to see the monomer crossover, as indicated on the plot. The single orientation curve is also shown to give one an idea of the importance of the orientational averaging in the appearance of

these power-law features in the intensity curves.

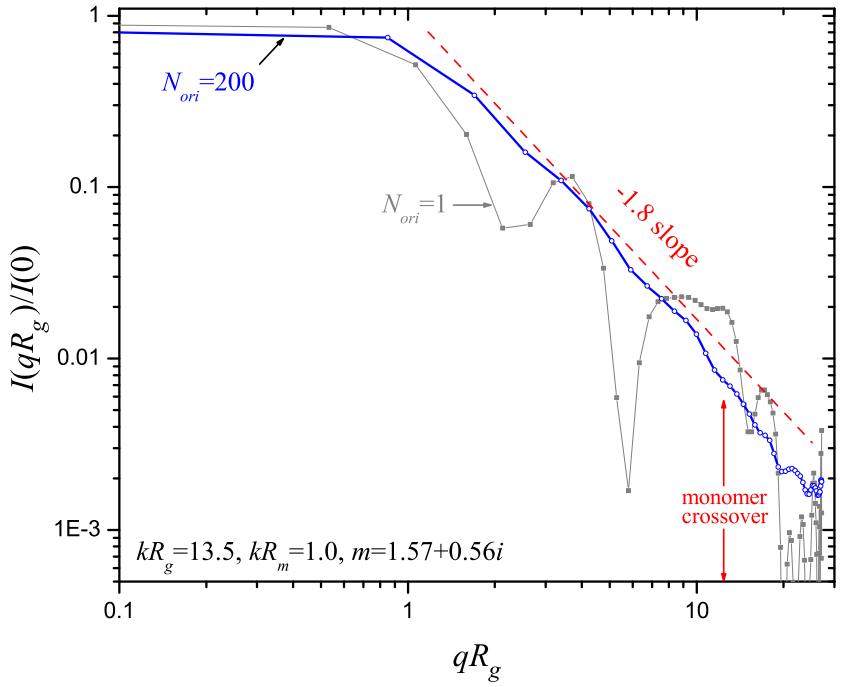

Figure 8.9: Normalized scattered intensity for the same aggregate as in the figures above, except here the monomer size is increased to $kR_m = 1.0$, which displays the monomer crossover.

8.3.2 Hybrid Superaggregates

Work by [88] shows that in some circumstances a system of growing DLCA aggregates can result in aggregates of aggregates, called superaggregates. A unique feature of a superaggregate is that while the component aggregates have a fractal dimension of $D_f \simeq 1.8$, the superaggregate, overall, has a dimension of $D_f \simeq 2.6$. Therefore, if the length scale interpretation of the crossovers in q-space analysis is valid, this superaggregate hybrid morphology should be apparent as a crossover in the power-law region of the scattering curves. The simulated superaggregate[5] used in the following calculations has $kR_g = 6.21$ and $N_m = 4275$ and is shown in plot (a) of Fig. (8.10) and Fig. (8.11).

The difficulty, however, in applying the DDA to a superaggregate is that they are much

[5]This superaggregate was provided courtesy of Flint Pierce [84].

larger than the aggregates considered thus far. Because orientational averaging must still be done, the computational requirements of the DDA calculations exceeds what is currently available in desktop computers. Consequently, a parallelized version of the DDA was developed and run on the supercomputers at the NASA High-End Computing (NASA HEC) center. Even with this formidable computational resource, the following DDA-based superaggregate calculations severely suffer from poor lattice fineness and orientational averaging.

To gain some insight regarding the effect of DDA lattice fineness on the resulting scattering quantities, Fig. (8.10) shows the single orientation and orientational ensemble averaged normalized q-space scattering curves for the superaggregate shown with two different DDA lattices. In plot (a), the lattice is very coarse, with only one site (dipole) per monomer, i.e. $N_{dda}/N_m = 1$. Plot (b) shows a finer lattice with $N_{dda}/N_m = 116$. Notice that as the lattice fineness increases, the number of orientations N_{ori} in the ensemble average decreases; this is, again, due to the limitations of the computational resources available. One can see in these plots that two power-law trends are indeed seen. This crossover seems to indicate the presence of two fractal dimensions in the superaggregate. The crossover occurs roughly at $qR_g \simeq 3$, i.e., at a length scale corresponding to roughly one-third of the R_g of the superaggregate. Close inspection of the image of the superaggregate shown in plot (a), and Fig. (8.11) below, does vaguely display a cluster-like appearance at this length scale. Moreover, since the $D_f \simeq 1.8$ length scale is expected to be smaller than the overall superaggregate size, one would expect to see this fractal dimension appear at high q values in the scattering curve[6]. The problem, however, is that the fractal dimension at smaller q, which should be representative of the larger length scale morphology, is roughly twice the expected value of $D_f \simeq 2.6$. Judging from the monomer size parameter $kR_m = 0.21$ and the fineness of the finest lattice [plot (b)], both of which are good values, the failure of the calculation to show the $D_f \simeq 2.6$ power-law is probably due to the poor orientational averaging, which includes only $N_{ori} = 75$ orientations.

[6]Remember the inverse relationship between q and R.

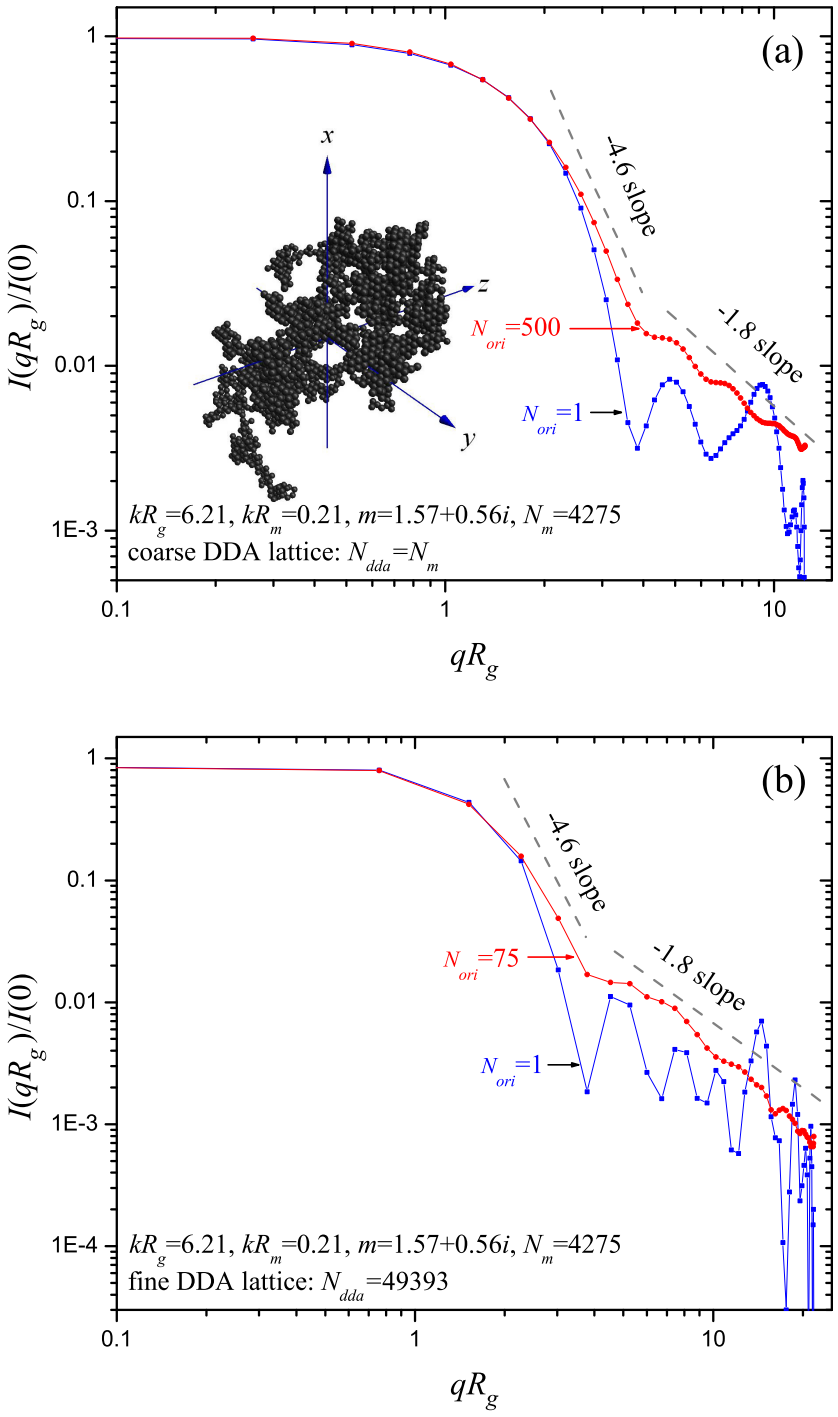

Figure 8.10: Normalized scattered intensity for a superaggregate. Plot (a) shows the fixed and orientationally averaged curve when a coarse DDA lattice is used to represent the aggregate; $N_{dda}/N_m = 1$, i.e., one lattice site per monomer. Plot (b) shows the same except for a finer DDA lattice of $N_{dda}/N_m = 116$. Notice that fewer orientations are used in plot (b) as a consequence of the limited computational resources.

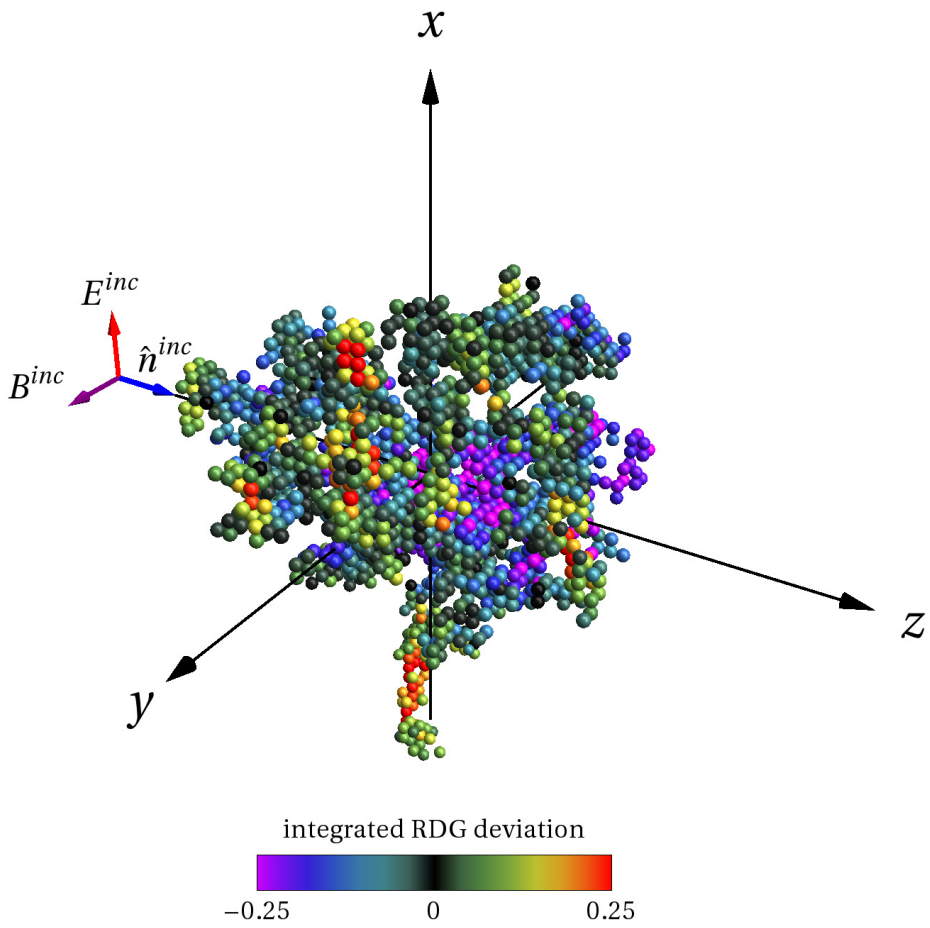

Figure 8.11: Same as Fig. (8.3) except for the superaggregate in Fig. (8.10).

As a final consideration, Figs. (8.11) and (8.12) show the plots analogous to Figs. (8.3) and plot (a) of Fig. (8.4) for the superaggregate. These plots show rather large deviations of the internal field from the RDG expectation. Overall, one can see that the RDG approximation appears to be less valid for the superaggregate as compared to the single aggregates above. The distribution in Fig. (8.12) is broader than the corresponding single aggregate deviation in (a) of Fig. (8.4). This behavior is somewhat expected since the superaggregate is more compact overall as compared its constituent aggregate; this causes stronger coupling between the monomers, resulting in greater deviations from the RDG expectation.

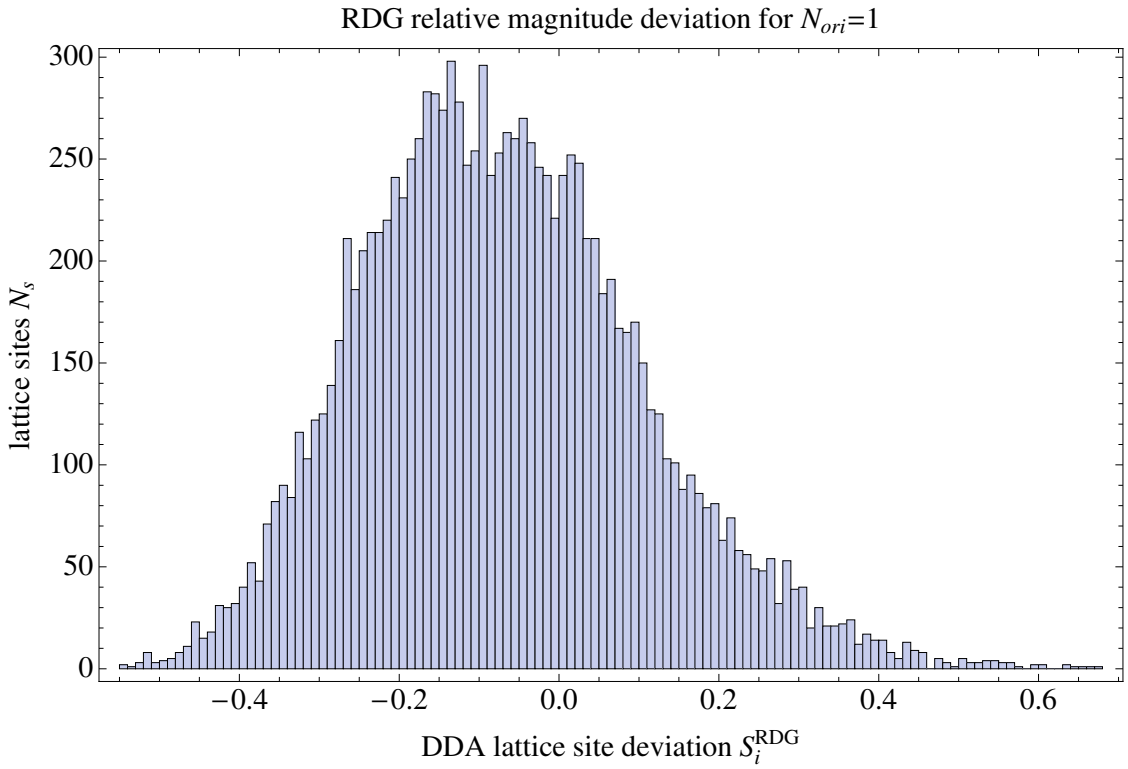

Figure 8.12: Same as plot (a) of Fig. (8.4) except for the superaggregate in Figs. (8.11) and (8.10).

Chapter 9

Diffraction

"The effect is a general characteristic of wave phenomena occurring whenever a portion of a wavefront, be it sound, a matter wave, or light, is obstructed in some way."

-E. Hecht [59, p. 443]

The diffraction of a wave by an aperture in an planar opaque screen is one the oldest problems in classical electromagnetic theory. The typical way to describe diffraction is via Huygens' principle, which gives the far-field wave behind the aperture in terms of a superposition of fictitious wavelet sources spanning the aperture opening. Although the Huygens' principle explanation is mathematically simple and accurate for regions where one is usually interested in its effects, i.e., the far-field, it is ultimately a misleading conceptual model for the physical process responsible for diffraction.

A more realistic picture of diffraction is depicted in Fig. (9.1). The incident wave induces surface current on the screen, which radiate to account for the observed reflected and transmitted wave at $\mathbf{r}$. In this sense the diffraction effect is manifest from the superposition of the wavelet fields radiated by the induced currents on the screen and the lack of any such current in the aperture. More interesting, perhaps, is that this surface current must somehow cancel the incident wave in regions behind the screen where no wave is observed. The objective in the following is to demonstrate that diffraction can be rigorously described within the framework of the microphysical model. In the process, the familiar Huygens' principle will be shown to be contained in the microphysical description as a consequence of

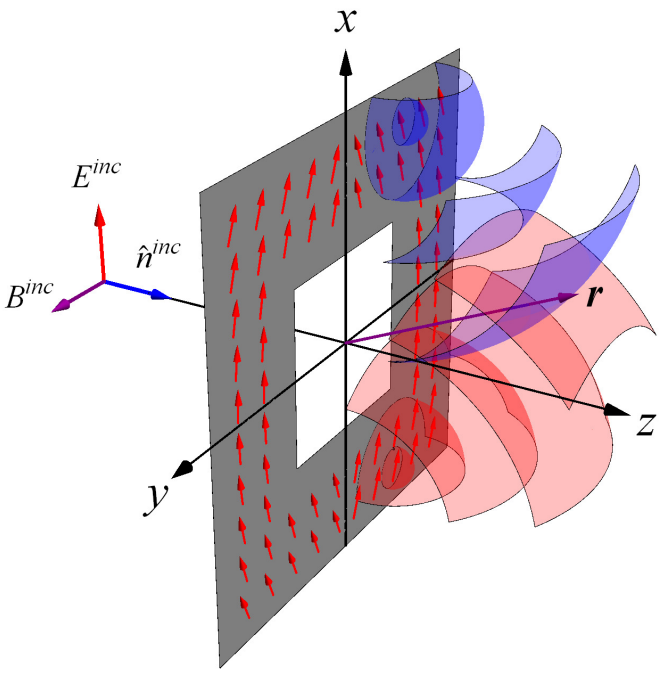

Figure 9.1: Microphysical model of planar diffraction. The red arrows on the screen represent induced surface current that radiate wavelets, shown by the spherical surfaces, to the observation point.

Babinet's principle and the far-field limit. Inherent to this microphysical description will be the Ewald-Oseen extinction theorem, which is not contained in the simple Huygens' picture, and will be responsible for cancellation of the incident wave in the dark regions behind the screen. The DDA will be used to simulate diffraction from an arbitrarily shaped aperture and find the induced current on the aperture's screen.

9.1 Infinite Plane

Before confronting the full treatment of an arbitrary aperture, it will be useful to consider the case of no aperture, i.e., an infinitesimally thin complete opaque plate. Recall that this situation is briefly considered in Sec. 2.3.2, where it is shown that a planar current sheet radiates a transverse plane wave. The calculations leading to that result are based on the microphysical model in that the total field is described in terms of the superposition of the wavelets radiated by the surface current elements rather than by matching boundary con-

ditions. Here the latter route is taken, which in comparison to the microphysical approach of Sec. 2.3.2 will demonstrate the utility of the microphysical perspective in understanding why the radiated wave takes the form that it does.

Consider the arrangement presented in Fig. (9.2). Here the plate lies in the x-y plane passing through the origin. Illuminating the plate is the plane wave of Eqs. (3.1) and (3.2), which strikes the plate normal to its surface and is polarized parallel to it, along the x-axis. The meaning of "opaque plate" here is vague. Opacity represents the material properties of the plate that are required in order to result in no transmission of the incident wave. This can be achieved, for example, by a perfectly conducting (PC) material, or by a perfectly absorbing (PA) material. Because of certain difficulties[1] associated with the theoretical description of PA materials, see [47, p. 183], the plate will be taken to be a perfect conductor throughout this work.

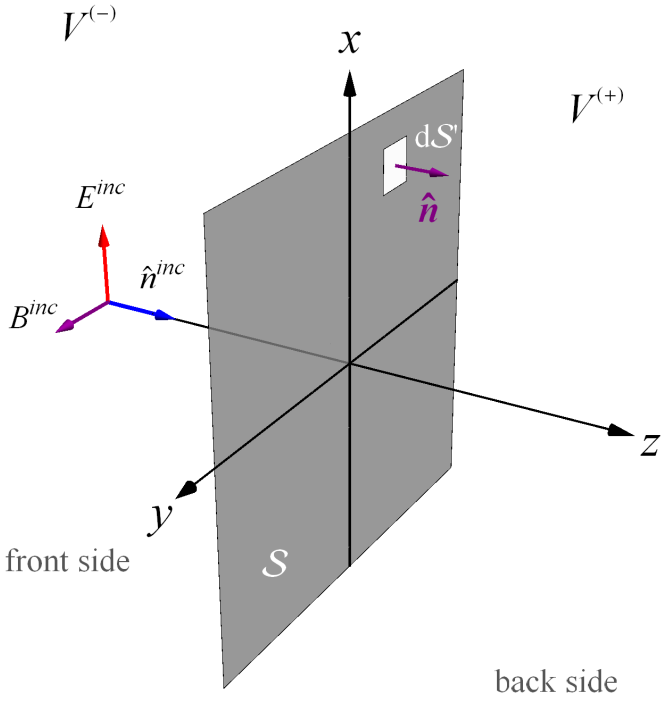

Figure 9.2: Geometry and notation for the perfectly conducting infinite plate.

The sketch in Fig. (9.2) shows that the plate separates space into two regions; $V^{(-)}$

[1]Basically, the plate cannot be both perfectly absorbing and infinitesimally thin.

denotes all space in front of the plate, and $V^{(+)}$ behind the plate, i.e.,

$$\mathbf{r} \in \begin{cases} V^{(-)} & z < 0 \\ V^{(+)} & z \geqslant 0. \end{cases} \tag{9.1}$$

The surface and surface normal for the plate are denoted by $\mathcal{S}$ and $\hat{\mathbf{n}}$, respectively. It will be assumed that the sources that produce the incident wave of Eqs. (3.1) and (3.2) are at infinity in $V^{(-)}$.

Because the plate is perfectly conducting, there is no wave behind the plate in $V^{(+)}$. The incident wave is reflected from the front side in $V^{(-)}$ producing a standing wave there. This recognition can be combined with the boundary conditions of Eqs. (1.40)-(1.43) to quickly arrive at the total fields on each side of the plate. Let the field on either side of the plate be denoted $\mathbf{E}^{(-)}$ and $\mathbf{E}^{(+)}$. Then, the boundary conditions give

$$\left.\begin{array}{l} \hat{\mathbf{n}} \times \mathbf{E}^{(+)}(\mathbf{r}) - \hat{\mathbf{n}} \times \mathbf{E}^{(-)}(\mathbf{r}) = 0 \\ \hat{\mathbf{n}} \times \mathbf{B}^{(+)}(\mathbf{r}) - \hat{\mathbf{n}} \times \mathbf{B}^{(-)}(\mathbf{r}) = \mu_o \mathbf{K}(\mathbf{r}) \end{array}\right\} \mathbf{r} \in \mathcal{S}. \tag{9.2}$$

The total fields are zero $\mathbf{E}^{(+)} = \mathbf{B}^{(+)} = 0$ for points behind the plate $\mathbf{r} \in V^{(+)}$. In front of the plate there is the superposition of the incident and reflected waves. Let the reflected wave be given by

$$\mathbf{E}^{ref}(\mathbf{r}) = -E_o^{inc} \exp(-ikr\hat{\mathbf{n}}^{inc} \cdot \hat{\mathbf{r}})\, \hat{\mathbf{x}} \quad \mathbf{r} \in V^{(-)}, \tag{9.3}$$

$$\mathbf{B}^{ref}(\mathbf{r}) = \frac{k}{\omega} E_o^{inc} \exp(-ikr\hat{\mathbf{n}}^{inc} \cdot \hat{\mathbf{r}})\, \hat{\mathbf{y}} \quad \mathbf{r} \in V^{(-)}, \tag{9.4}$$

and the wave behind the plate

$$\mathbf{E}^{(+)}(\mathbf{r}) = 0 \quad \mathbf{r} \in V^{(+)}, \tag{9.5}$$

$$\mathbf{B}^{(+)}(\mathbf{r}) = 0 \quad \mathbf{r} \in V^{(+)}. \tag{9.6}$$

Then from Eqs. (3.1), (3.2), and (9.2), one finds that the surface current that must reside on the plate to produce this field configuration is

$$\mathbf{K}(\mathbf{r}) = \frac{2k}{\omega\mu_o} E_o^{inc}\, \hat{\mathbf{x}}, \quad \mathbf{r} \in \mathcal{S}. \tag{9.7}$$

The total fields on either side of the plate are then

$$\mathbf{E}(\mathbf{r}) = \begin{cases} 2iE_o^{inc} \sin\left(k\mathbf{r}\cdot\hat{\mathbf{n}}^{inc}\right)\hat{\mathbf{x}} & \mathbf{r}\in V^{(-)} \\ 0 & \mathbf{r}\in V^{(+)}, \end{cases} \tag{9.8}$$

$$\mathbf{B}(\mathbf{r}) = \begin{cases} 2\frac{k}{\omega}E_o^{inc} \cos\left(k\mathbf{r}\cdot\hat{\mathbf{n}}^{inc}\right)\hat{\mathbf{x}} & \mathbf{r}\in V^{(-)} \\ 0 & \mathbf{r}\in V^{(+)}. \end{cases} \tag{9.9}$$

One can see that the total fields in Eqs. (9.8) and (9.9) constitute a transverse standing wave in $V^{(-)}$ and zero wave in $V^{(+)}$ as expected. This result is essentially the same as the far more difficult calculations of Sec. 2.3.2, except here it is basically deduced from a physical expectation of the form of the wave on either side of the plate in combination with the boundary conditions.

The total wave represented as the superposition of the indent wave and scattered waves

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \mathbf{E}^{sca}(\mathbf{r}),$$

shows that

$$\mathbf{E}^{sca}(\mathbf{r}) = \begin{cases} -\left[\mathbf{E}^{inc}(\mathbf{r})\right]^* & \mathbf{r}\in V^{(-)} \\ -\mathbf{E}^{inc}(\mathbf{r}) & \mathbf{r}\in V^{(+)}. \end{cases} \tag{9.10}$$

In addition, by comparing the surface current in Eq. (2.52) to (9.7), one finds that

$$\mathbf{E}^{sca}(\mathbf{r}) = -\mathbf{E}^{inc}(\mathbf{r}) = -\mathbf{E}^{pl}(\mathbf{r}) \quad \mathbf{r}\in V^{(+)}.$$

When combined with Eq. (2.43), this reveals that the induced surface current of Eq. (9.7) radiates a wave that travels away from the plate with the same magnitude as the incident wave but shifted by π in phase. In front of the plate, the incident and scattered (radiated) waves are counter-propagating and hence form a standing wave with a resulting magnitude of twice either of the components waves. Behind the plate, however, the waves are co-propagating and the phase shift causes them to cancel. One can then understand that it is the *radiation from the induced surface current that produces a wave that cancels the incident wave behind the plate*, resulting in no wave there. Therefore, the observation of a diffracted wave behind the plate when an aperture is present will be seen to be a consequence of the incomplete cancellation caused by the absence of induced current in the aperture; it is not strictly an (aperture) "edge effect."

9.2 Arbitrary Apertures

Just as Eq. (2.26) is the fundamental relationship between a three-dimensional object's induced current density $\mathbf{J}$ and its scattered field, Eq. (2.43) is the analogous relation for two-dimensional surface currents $\mathbf{K}$ and its corresponding scattered field. Then,

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + i\omega\mu_o \int_{\mathcal{S}} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}(\mathbf{r}')\, \mathrm{d}\mathcal{S}', \tag{9.11}$$

is the complete solution to the Maxwell equations for the fields on either side of $\mathcal{S}$. One can readily derive the corresponding expression for the magnetic field following the last half of Sec. 2.1.

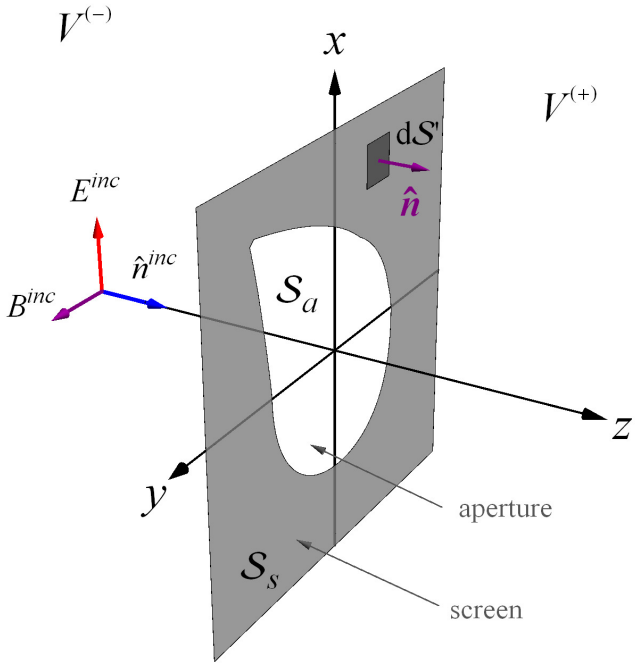

Figure 9.3: Sketch like Fig. (9.2) except here an arbitrarily shaped aperture is present in the screen.

There is a subtle issue associated with Eq. (9.11). The integral solution in the Green's function method relies on an application of the second vector-dyadic Green's theorem. This integral theorem is implicitly used in Sec. 2.1 to derive the VIEs but is not explicitly stated there for the sake of brevity. The validity of the integral theorem relies on the continuity

of the fields residing on either side of the integration surface. Since the aperture in the infinitesimally thin screen creates such an edge, the fields at the aperture edge become infinite [89, Sec. 1.5]. If the fields are also discontinuous across the contour defining the edge, the integral theorem could become invalid. It turns out that although the fields are infinite at the aperture edge, they are also continuous across the edge contour [89, p. 25]. The second vector-dyadic Green's theorem is then not violated by the discontinuity in the surface current at the aperture. The following will derive Eq. (9.11) directly from the Maxwell equations using Green's integral theorem.

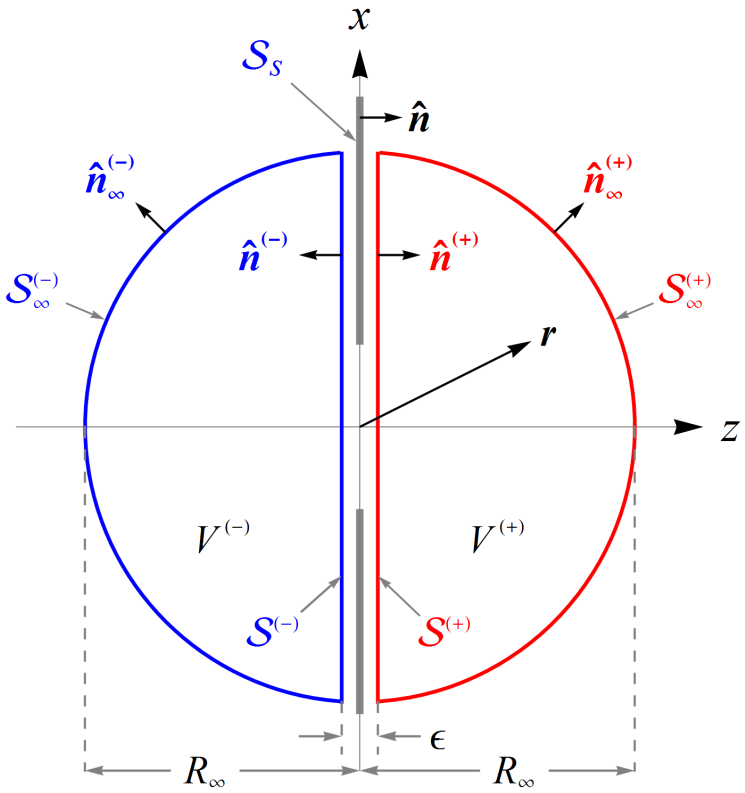

Figure 9.4: Division of space surrounding the screen.

Let the planar surfaces representing an aperture and the screen be denoted by $\mathcal{S}_a$ and $\mathcal{S}_s$, respectively, see Fig. (9.3). The union of the screen and aperture surfaces makes the complete planar surface of Fig. (9.2), i.e., $\mathcal{S}_s \cup \mathcal{S}_a = \mathcal{S}$. Now let the half-spaces on either side of the screen be defined by the volumes $V^{(-)}$ and $V^{(+)}$ enclosed by the closed hemispherical-cap surfaces shown in Fig. (9.4). These hemispherical-cap surfaces are formed

by the joining together of circular planar surfaces $\mathcal{S}^{(-)}$ and $\mathcal{S}^{(+)}$ with the open hemispherical surfaces $\mathcal{S}_\infty^{(-)}$ and $\mathcal{S}_\infty^{(+)}$ of radius R_∞. The surface normals $\hat{\mathbf{n}}_\infty^{(-)}$ and $\hat{\mathbf{n}}_\infty^{(+)}$ to $\mathcal{S}_\infty^{(-)}$ and $\mathcal{S}_\infty^{(+)}$, respectively, point outward from the volume enclosed. The surface normals to the circular planar segments $\hat{\mathbf{n}}^{(-)}$ and $\hat{\mathbf{n}}^{(+)}$ point *inward* into the volumes $V^{(-)}$ and $V^{(+)}$, respectively. The circular planar surfaces $\mathcal{S}^{(-)}$ and $\mathcal{S}^{(+)}$ are oriented parallel to the screen surface $\mathcal{S}_s$ and separated from it by a distance $\epsilon/2$. The observation point $\mathbf{r}$ will be restricted in the following to lie in either $V^{(-)}$ or $V^{(+)}$.

Following the calculations given in Sec. 1.3, the Maxwell equations for points in $V^{(-)}$ or $V^{(+)}$ can be combined into a single wave equation

$$\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{E}(\mathbf{r}) - k^2 \mathbf{E}(\mathbf{r}) = 0 \qquad \mathbf{r} \in V^{(-)} \vee V^{(+)}. \tag{9.12}$$

Recall that $\mathbf{r}$ is excluded from $\mathcal{S}_s$ since the surfaces enclosing $V^{(-)}$ and $V^{(+)}$ are lifted off $\mathcal{S}_s$ by $\epsilon/2$, see Fig. (9.4).

Next consider the second vector-dyadic Green's theorem [90]

$$\int_V \Big\{ \mathbf{A}(\mathbf{r}) \cdot \big[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{C}}(\mathbf{r}) \big] - \overleftrightarrow{\mathbf{C}}(\mathbf{r}) \cdot \big[\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r}) \big] \Big\} \, \mathrm{d}\mathbf{r}$$
$$= \oint_{\mathcal{S}} \hat{\mathbf{n}} \cdot \Big\{ \overleftrightarrow{\mathbf{C}}(\mathbf{r}) \times \big[\boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r}) \big] - \mathbf{A}(\mathbf{r}) \times \big[\boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{C}}(\mathbf{r}) \big] \Big\} \, \mathrm{d}\mathcal{S}, \tag{9.13}$$

where $\mathcal{S}$ is a the surface bounding V and $\hat{\mathbf{n}}$ is the outward surface normal. Now make the substitution

$$\mathbf{A}(\mathbf{r}) = \mathbf{E}(\mathbf{r}), \quad \overleftrightarrow{\mathbf{C}}(\mathbf{r}) = \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}'),$$

in Eq. (9.13) and apply the integral theorem to all space $\mathbb{R} = V^{(-)} \cup V^{(+)}$, which is bounded by the hemispherical cap surfaces in the limit that $\epsilon \to 0$ and $R_\infty \to \infty$. This is where the tricky point appears. The electric field is infinite at the aperture edge and this singularity is included in the parts of the total surface integral that cover the planar surfaces $\mathcal{S}^{(-)}$ and $\mathcal{S}^{(+)}$ in the limit $\epsilon \to 0$. Moreover, the dyadic Green's function explodes if the observation point ever coincides with the integration point. This is why the space surrounding the screen is partitioned in the bizarre way shown in Fig. (9.4). Since there are two volumes $V^{(-)}$ and

$V^{(+)}$, and four surfaces $\mathcal{S}^{(-)}$, $\mathcal{S}^{(+)}$, $\mathcal{S}_\infty^{(-)}$, and $\mathcal{S}_\infty^{(+)}$, the above substitution results in the following integral relation

$$\begin{aligned}
&\int_{V^{(-)}\cup V^{(+)}} \Big\{\mathbf{E}(\mathbf{r})\cdot\big[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]-\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\cdot\big[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{E}(\mathbf{r})\big]\Big\}\,\mathrm{d}\mathbf{r}\\
&\qquad=\int_{\mathcal{S}_\infty^{(-)}}\hat{\mathbf{n}}_\infty^{(-)}\cdot\Big\{\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\big[\boldsymbol{\nabla}\times\mathbf{E}(\mathbf{r})\big]-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad-\int_{\mathcal{S}^{(-)}}\hat{\mathbf{n}}^{(-)}\cdot\Big\{\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\big[\boldsymbol{\nabla}\times\mathbf{E}(\mathbf{r})\big]-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad-\int_{\mathcal{S}^{(+)}}\hat{\mathbf{n}}^{(+)}\cdot\Big\{\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\big[\boldsymbol{\nabla}\times\mathbf{E}(\mathbf{r})\big]-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad+\int_{\mathcal{S}_\infty^{(+)}}\hat{\mathbf{n}}_\infty^{(+)}\cdot\Big\{\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\big[\boldsymbol{\nabla}\times\mathbf{E}(\mathbf{r})\big]-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad\qquad\qquad\qquad\mathbf{r}\notin\mathcal{S}_s\;\epsilon\to 0,\;R_\infty\to\infty,\quad(9.14)
\end{aligned}$$

Notice that since the surface normals $\hat{\mathbf{n}}^{(-)}$ and $\hat{\mathbf{n}}^{(+)}$ point into the volumes $V^{(-)}$ and $V^{(+)}$, they bring in a negative sign in two of the last integrals in Eq. (9.14). Next, using Eqs. (1.22), (2.2), and (1.35), these integrals become, after some algebra,

$$\begin{aligned}
&\int_{V^{(-)}\cup V^{(+)}}\mathbf{E}(\mathbf{r})\delta\left(\mathbf{r}-\mathbf{r}'\right)\,\mathrm{d}\mathbf{r}\\
&\qquad=\int_{\mathcal{S}_\infty^{(-)}}\hat{\mathbf{n}}_\infty^{(-)}\cdot\Big\{i\omega\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\mathbf{B}(\mathbf{r})-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad-\int_{\mathcal{S}^{(-)}}\hat{\mathbf{n}}^{(-)}\cdot\Big\{i\omega\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\mathbf{B}(\mathbf{r})-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad-\int_{\mathcal{S}^{(+)}}\hat{\mathbf{n}}^{(+)}\cdot\Big\{i\omega\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\mathbf{B}(\mathbf{r})-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad+\int_{\mathcal{S}_\infty^{(+)}}\hat{\mathbf{n}}_\infty^{(+)}\cdot\Big\{i\omega\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\times\mathbf{B}(\mathbf{r})-\mathbf{E}(\mathbf{r})\times\big[\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')\big]\Big\}\,\mathrm{d}\mathcal{S}\\
&\qquad\qquad\qquad\qquad\mathbf{r}\notin\mathcal{S}_s\;\epsilon\to 0,\;R_\infty\to\infty.\quad(9.15)
\end{aligned}$$

Now assume that the infinite sheet has a limited but very large extent. Then, since the radius R_∞ of the hemispherical surfaces $\mathcal{S}_\infty^{(-)}$ and $\mathcal{S}_\infty^{(+)}$ goes to infinity, the vector and dyadic fields appearing in Eq. (9.15) will vanish due to the Sommerfeld radiation conditions

$$\lim_{R_\infty\to\infty}\left\{R_\infty\left[\boldsymbol{\nabla}\times\mathbf{E}(R_\infty\hat{\mathbf{r}})-ik\hat{\mathbf{r}}\times\mathbf{E}(R_\infty\hat{\mathbf{r}})\right]\right\}=0,\tag{9.16}$$

$$\lim_{R_\infty \to \infty} \left\{ R_\infty \left[\boldsymbol{\nabla} \times \mathbf{B}(R_\infty \hat{\mathbf{r}}) - ik\hat{\mathbf{r}} \times \mathbf{B}(R_\infty \hat{\mathbf{r}}) \right] \right\} = 0, \tag{9.17}$$

$$\lim_{R_\infty \to \infty} \left\{ R_\infty \left[\boldsymbol{\nabla} \times \overleftrightarrow{\mathbf{G}}_e(R_\infty \hat{\mathbf{r}}, \mathbf{r}') - ik\hat{\mathbf{r}} \times \overleftrightarrow{\mathbf{G}}_e(R_\infty \hat{\mathbf{r}}, \mathbf{r}') \right] \right\} = 0. \tag{9.18}$$

Combining Eq. (9.15) with (9.16)-(9.18) and interchanging $\mathbf{r}$ and $\mathbf{r}'$, which means that $\boldsymbol{\nabla} \to \boldsymbol{\nabla}'$, one finds

$$\begin{aligned} &\int_{V^{(-)} \cup V^{(+)}} \mathbf{E}(\mathbf{r}') \delta(\mathbf{r}' - \mathbf{r}) \, \mathrm{d}\mathbf{r}' \\ &\quad = -\int_{\mathcal{S}^{(-)}} \hat{\mathbf{n}}^{(-)} \cdot \left\{ i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \times \mathbf{B}(\mathbf{r}') - \mathbf{E}(\mathbf{r}') \times \left[\boldsymbol{\nabla}' \times \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \right] \right\} \mathrm{d}\mathcal{S}' \\ &\quad\quad - \int_{\mathcal{S}^{(+)}} \hat{\mathbf{n}}^{(+)} \cdot \left\{ i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \times \mathbf{B}(\mathbf{r}') - \mathbf{E}(\mathbf{r}') \times \left[\boldsymbol{\nabla}' \times \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \right] \right\} \mathrm{d}\mathcal{S}' \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mathbf{r} \notin \mathcal{S}_s \; \epsilon \to 0. \end{aligned} \tag{9.19}$$

Evaluating the integral on the right-hand side and using Eq. (2.13), one finds

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= -\int_{\mathcal{S}^{(-)}} \hat{\mathbf{n}}^{(-)} \cdot \left[i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \times \mathbf{B}(\mathbf{r}') - \mathbf{E}(\mathbf{r}') \times \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}, \mathbf{r}') \right] \mathrm{d}\mathcal{S}' \\ &\quad - \int_{\mathcal{S}^{(+)}} \hat{\mathbf{n}}^{(+)} \cdot \left[i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \times \mathbf{B}(\mathbf{r}') - \mathbf{E}(\mathbf{r}') \times \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}, \mathbf{r}') \right] \mathrm{d}\mathcal{S}' \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mathbf{r} \notin \mathcal{S}_s \; \epsilon \to 0. \end{aligned} \tag{9.20}$$

Next, one can show that

$$\hat{\mathbf{n}} \cdot \left[\overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \times \mathbf{B}(\mathbf{r}') \right] = -\overleftrightarrow{\mathbf{G}}_e(\mathbf{r}, \mathbf{r}') \cdot \left[\hat{\mathbf{n}} \times \mathbf{B}(\mathbf{r}') \right],$$

$$\hat{\mathbf{n}} \cdot \left[\mathbf{E}(\mathbf{r}') \times \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}', \mathbf{r}) \right] = -\overleftrightarrow{\mathbf{G}}_m(\mathbf{r}', \mathbf{r}) \cdot \left[\hat{\mathbf{n}} \times \mathbf{E}(\mathbf{r}') \right],$$

so that Eq. (9.20) becomes

$$\begin{aligned} \mathbf{E}(\mathbf{r}) &= \int_{\mathcal{S}^{(-)}} \left\{ i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \cdot \left[\hat{\mathbf{n}}^{(-)} \times \mathbf{B}(\mathbf{r}') \right] - \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}', \mathbf{r}) \cdot \left[\hat{\mathbf{n}}^{(-)} \times \mathbf{E}(\mathbf{r}') \right] \right\} \mathrm{d}\mathcal{S}' \\ &\quad + \int_{\mathcal{S}^{(+)}} \left\{ i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r}', \mathbf{r}) \cdot \left[\hat{\mathbf{n}}^{(+)} \times \mathbf{B}(\mathbf{r}') \right] - \overleftrightarrow{\mathbf{G}}_m(\mathbf{r}', \mathbf{r}) \cdot \left[\hat{\mathbf{n}}^{(+)} \times \mathbf{E}(\mathbf{r}') \right] \right\} \mathrm{d}\mathcal{S}' \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \mathbf{r} \notin \mathcal{S}_s \; \epsilon \to 0. \end{aligned} \tag{9.21}$$

In the limit that $\epsilon \to 0$, the surfaces $\mathcal{S}^{(-)}$ and $\mathcal{S}^{(+)}$ snap onto the planar surface formed by the union of the sheet and aperture $\mathcal{S}_s$ and $\mathcal{S}_a$, respectively. Taking this limit and noticing

that $-\hat{\mathbf{n}}^{(-)} = \hat{\mathbf{n}}^{(+)} = \hat{\mathbf{n}}$, recall Figs. (9.2) and (9.4), one finds that Eq. (9.23) becomes

$$\mathbf{E}(\mathbf{r}) = \int_{\mathcal{S}} \Big\{ i\omega \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \big[\hat{\mathbf{n}} \times \mathbf{B}^{(+)}(\mathbf{r}') - \hat{\mathbf{n}} \times \mathbf{B}^{(-)}(\mathbf{r}')\big] + \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \big[\hat{\mathbf{n}} \times \mathbf{E}^{(+)}(\mathbf{r}') - \hat{\mathbf{n}} \times \mathbf{E}^{(-)}(\mathbf{r}')\big] \Big\} \, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_s, \tag{9.22}$$

where the relations $\overleftrightarrow{\mathbf{G}}_e(\mathbf{r}',\mathbf{r}) = \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}')$ and $\overleftrightarrow{\mathbf{G}}_m(\mathbf{r}',\mathbf{r}) = -\overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}')$ have been used, and the notation $\mathbf{E}^{(-)}$ and $\mathbf{E}^{(+)}$ denotes the field values on $\mathcal{S}^{(-)}$ and $\mathcal{S}^{(+)}$, respectively. Using the boundary conditions of Eqs. (1.40) and (1.41), and noticing that there is no discontinuity in the incident fields across the screen, one finds that

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + i\omega\mu_o \int_{\mathcal{S}_s} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}(\mathbf{r}') \, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_s, \tag{9.23}$$

which is Eq. (9.11).

Equation (9.23) does not yet solve anything; it is an integral equation that gives the total electric field on either side of the screen in terms of the surface current on the screen. The surface current must still somehow be found, and this will be the point behind the application of the DDA later. The problem, however, is that the DDA is a numerical technique and as such cannot be directly applied to Eq. (9.23) since the screen is infinite in extent. This issue is resolved in the next section using Babinet's principle.

9.3 Babinet's Principle

The reflection symmetry of the planar diffraction problem, the perfect conductor limit of the screen material, and the duality of the Maxwell equations yield the Babinet principle [6, 91]. This principle allows an exact transformation of the infinite-in-extent screen problem into a finite-in-extent complimentary problem, the latter of which can be treated with the DDA. The following will derive the full form of Babinet's principle. Note that this is largely based on [91, Sec. 4.4].

Consider the screen shown in Fig. (9.2). This configuration will be called the original problem, or configuration (a), and is shown again in (a) of Fig. (9.5). The complimentary

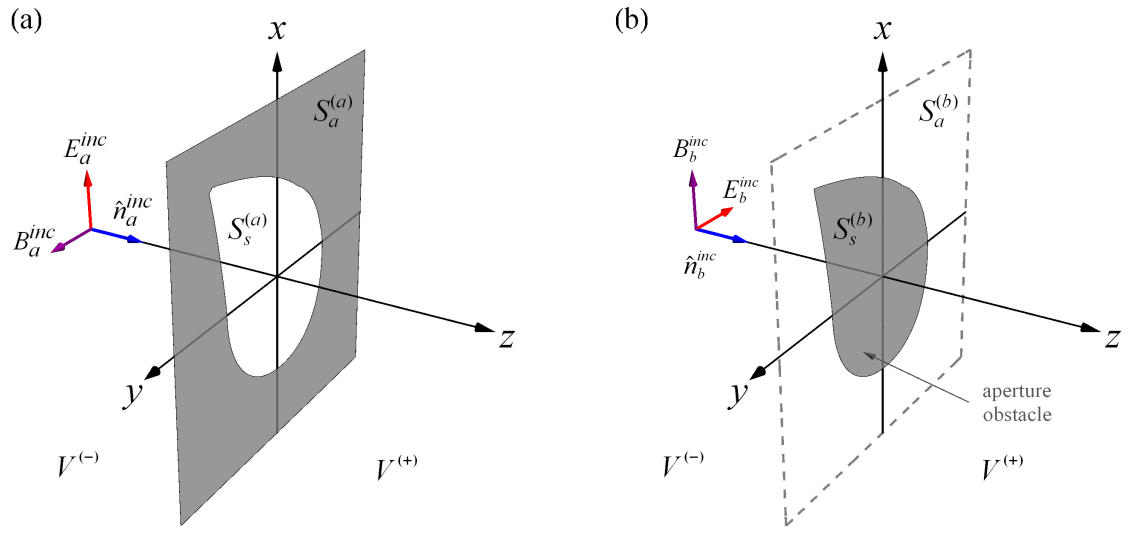

Figure 9.5: The original problem (a), and the complimentary problem (b), used in the formulation of Babinet's principle for locations on the *back side* of the screen.

configuration, also called configuration (b), is shown in (b) of Fig. (9.5) and is formed by removing the screen and making the aperture a screen, or aperture-obstacle. Notice in this latter configuration that the polarization of the incident wave is rotated by $-\pi/2$ in the right-handed sense. The notation $\mathcal{S}_a^{(a)}$ and $\mathcal{S}_s^{(a)}$ denotes the surfaces for the aperture and screen in the original problem, i.e. Fig (9.5) sketch (a). Likewise for $\mathcal{S}_a^{(b)}$ and $\mathcal{S}_s^{(b)}$ in the complimentary problem, sketch (b) of Fig (9.5). For the configuration in Fig (9.5) sketch (a), let

$$\mathbf{E}_a(\mathbf{r}) = \mathbf{E}_a^{inc}(\mathbf{r}) + \mathbf{E}_a^{sca}(\mathbf{r}), \tag{9.24}$$

$$\mathbf{B}_a(\mathbf{r}) = \mathbf{B}_a^{inc}(\mathbf{r}) + \mathbf{B}_a^{sca}(\mathbf{r}), \tag{9.25}$$

where the notation $\mathbf{E}_a$ denotes the field quantity for configuration a. From Eqs. (1.40) and (1.41), the boundary conditions at the screen surface give

$$\hat{\mathbf{n}} \times \mathbf{E}_a(\mathbf{r}) = 0 \quad \mathbf{r} \in \mathcal{S}_s^{(a)}, \tag{9.26}$$

$$\hat{\mathbf{n}} \times \mathbf{B}_a(\mathbf{r}) = \hat{\mathbf{n}} \times \mathbf{B}_a^{inc}(\mathbf{r}) \quad \mathbf{r} \in \mathcal{S}_a^{(a)}, \tag{9.27}$$

where Eq. (9.26) is the PC boundary condition, whereas Eq. (9.27) is shown by [91, p. 280]

to be due to symmetry.

Next consider the complimentary problem, sketch (b) in Fig. (9.5). Let

$$\mathbf{E}_b(\mathbf{r}) = \mathbf{E}_b^{inc}(\mathbf{r}) + \mathbf{E}_b^{sca}(\mathbf{r}) \tag{9.28}$$

$$\mathbf{B}_b(\mathbf{r}) = \mathbf{B}_b^{inc}(\mathbf{r}) + \mathbf{B}_b^{sca}(\mathbf{r}), \tag{9.29}$$

and the boundary conditions give

$$\hat{\mathbf{n}} \times \mathbf{E}_b(\mathbf{r}) = 0 \quad \mathbf{r} \in \mathcal{S}_a^{(b)} \tag{9.30}$$

$$\hat{\mathbf{n}} \times \mathbf{B}_b(\mathbf{r}) = \hat{\mathbf{n}} \times \mathbf{B}_b^{inc}(\mathbf{r}) \quad \mathbf{r} \in \mathcal{S}_a^{(b)}. \tag{9.31}$$

By comparing the direction of the incident fields in (a) and (b) of Fig. (9.5) and using Eqs. (2.58) and (2.59), one finds the following relationship between the fields in configurations (a) and (b),

$$\mathbf{E}_b^{inc}(\mathbf{r}) = -\frac{\omega}{k}\mathbf{B}_a^{inc}(\mathbf{r}), \tag{9.32}$$

$$\mathbf{B}_b^{inc}(\mathbf{r}) = \frac{k}{\omega}\mathbf{E}_a^{inc}(\mathbf{r}). \tag{9.33}$$

Consider the following substitutions, based on the duality of the Maxwell equations,

$$\mathbf{E}_b^{sca}(\mathbf{r}) \leftrightarrow \frac{\omega}{k}\mathbf{B}_a(\mathbf{r}) \quad \text{as} \quad \mathcal{S}_a^{(b)} \leftrightarrow \mathcal{S}_a^{(a)} \quad \mathbf{r} \in V^{(+)} \tag{9.34}$$

$$\mathbf{B}_b^{sca}(\mathbf{r}) \leftrightarrow -\frac{k}{\omega}\mathbf{E}_a(\mathbf{r}) \quad \text{as} \quad \mathcal{S}_s^{(b)} \leftrightarrow \mathcal{S}_s^{(a)} \quad \mathbf{r} \in V^{(+)}, \tag{9.35}$$

where it is important to note that that observation point is restricted to $V^{(+)}$. The Maxwell curl equations that must be satisfied in $V^{(+)}$ by the fields of Eqs. (9.24), (9.25), (9.28), and (9.29) are

$$\boldsymbol{\nabla} \times \mathbf{E}_a(\mathbf{r}) = i\omega\mathbf{B}_a(\mathbf{r}), \tag{9.36}$$

$$\boldsymbol{\nabla} \times \mathbf{B}_a(\mathbf{r}) = -i\omega\mu_o\varepsilon_o\mathbf{E}_a(\mathbf{r}), \tag{9.37}$$

and

$$\boldsymbol{\nabla} \times \mathbf{E}_b(\mathbf{r}) = i\omega\mathbf{B}_b(\mathbf{r}), \tag{9.38}$$

$$\boldsymbol{\nabla} \times \mathbf{B}_b(\mathbf{r}) = -i\omega\mu_o\varepsilon_o\mathbf{E}_b(\mathbf{r}). \tag{9.39}$$

By making the substitutions in Eqs. (9.34) and (9.35) in Eqs. (9.36) and (9.37), one obtains

$$\boldsymbol{\nabla} \times \mathbf{B}_b^{sca}(\mathbf{r}) = -i\omega\mu_o\varepsilon_o \mathbf{E}_b^{sca}(\mathbf{r}) \quad \mathbf{r} \in V^{(+)}, \tag{9.40}$$

$$\boldsymbol{\nabla} \times \mathbf{E}_b^{sca}(\mathbf{r}) = i\omega \mathbf{B}_b^{sca}(\mathbf{r}) \quad \mathbf{r} \in V^{(+)}, \tag{9.41}$$

which are exactly the Maxwell equations satisfied by the scattered fields of the complimentary problem in (b) of Fig. (9.5). Moreover, using Eqs. (9.34) and (9.35) again, one can verify that the boundary conditions of configuration (a) are satisfied by the scattered fields of configuration (b) given that $\mathcal{S}_a^{(a)} \to \mathcal{S}_a^{(b)}$ and $\mathcal{S}_s^{(a)} \to \mathcal{S}_s^{(b)}$. This result is Babinet's principle; the scattered fields of configuration (b) are solutions to the Maxwell equations for the total fields of configuration (a). This means that one can solve the complimentary problem, then use the substitutions in Eqs. (9.34) and (9.35) to automatically get the solution for the original problem. This is nice since the complimentary problem involves a finite screen, the aperture obstacle, as opposed to the infinite screen in the original problem. This then permits use of numerical techniques like the DDA to find the induced surface current.

Combination of Eqs. (9.34) and (9.35) with the boundary condition of Eq. (1.40) and the symmetry relations of [91, p. 280] yields the induced surface current on the screen in configuration (a)

$$\mathbf{K}_a(\mathbf{r}) = \frac{2k}{\omega\mu_o}\hat{\mathbf{n}} \times \mathbf{E}_b^{sca}(\mathbf{r}) + \frac{2k}{\omega\mu_o}\mathbf{E}_a^{inc}(\mathbf{r}) \quad \mathbf{r} \in \mathcal{S}_s^{(a)} \tag{9.42}$$

The problem is now finally solved. The DDA can be applied to the complimentary problem (b) to find the scattered field, which can then be fed into Eq. (9.42) to obtain the current induced in the original screen in configuration (a). Once the current is known, it can be used in Eq. (9.23) to get the diffracted wave on either side of the screen in the original problem. Notice that Babinet's principle is an exact transformation between the original problem and its compliment. There are no approximations needed in its derivation and hence its use to solve the diffraction problem in combination with the DDA will not introduce errors into the final solution beyond those inherent to the DDA.

9.4 Reduction to Huygens Principle

It is now possible to show how the familiar Huygens principle is contained in Eq. (9.23). The scattered field due to the screen in Fig. (9.3) is given by this equation as

$$\mathbf{E}^{sca}(\mathbf{r}) = i\omega\mu_o \int_{\mathcal{S}_s} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}(\mathbf{r}')\, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_s. \tag{9.43}$$

Next, recall from Eq. (9.34) that the scattered field due to configuration (b) is related through Babinet's principle to the total field of configuration (a). Let the surface current induced on the screen in configuration (a), $\mathcal{S}_s^{(a)}$, be denoted by $\mathbf{K}_a$, and $\mathbf{K}_b$ for configuration (b), i.e. the aperture-obstacle current. Then Eq. (9.43) applied to these two currents sources read

$$\mathbf{E}_a^{sca}(\mathbf{r}) = i\omega\mu_o \int_{\mathcal{S}_s^{(a)}} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}_a(\mathbf{r}')\, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_s^{(a)}, \tag{9.44}$$

and

$$\mathbf{E}_b^{sca}(\mathbf{r}) = i\omega\mu_o \int_{\mathcal{S}_a^{(b)}} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}_b(\mathbf{r}')\, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_a^{(b)}. \tag{9.45}$$

But from Eq. (9.34), Eq. (9.45) is equivalent to $(\omega/k)\mathbf{B}_a$. Then,

$$\mathbf{B}_a(\mathbf{r}) = ik\mu_o \int_{\mathcal{S}_a^{(b)}} \overleftrightarrow{\mathbf{G}}_e(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}_b(\mathbf{r}')\, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_a^{(b)}. \tag{9.46}$$

The total electric field for configuration (a) follows from Eq. (9.46) from the Maxwell curl equation (1.22), giving

$$\mathbf{E}_a(\mathbf{r}) = -\frac{\omega\mu_o}{k} \int_{\mathcal{S}_a^{(b)}} \overleftrightarrow{\mathbf{G}}_m(\mathbf{r},\mathbf{r}') \cdot \mathbf{K}_b(\mathbf{r}')\, \mathrm{d}\mathcal{S}' \quad \mathbf{r} \notin \mathcal{S}_a^{(b)}, \tag{9.47}$$

where Eq. (2.13) has been used. Equations (9.46) and (9.47) are exact relations and demonstrate that the total fields for configuration (a) can be expressed in terms of the current induced on the screen in configuration (b).

Substantial simplification of Eq. (9.47) results if the far-field form of the dyadic Green's functions are used. The results are

$$\mathbf{E}_a(\mathbf{r}) = i\omega\mu_o \frac{\exp(ikr)}{4\pi r} \int_{\mathcal{S}_a^{(b)}} \hat{\mathbf{r}} \times \mathbf{K}_b(\mathbf{r}') \exp\left(-ik\hat{\mathbf{r}} \cdot \mathbf{r}'\right)\, \mathrm{d}\mathcal{S}' \quad kr \to \infty, \tag{9.48}$$

and

$$\mathbf{B}_a(\mathbf{r}) = ik\mu_o \frac{\exp(ikr)}{4\pi r} \int_{\mathcal{S}_a^{(b)}} \left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}} \right) \cdot \mathbf{K}_b(\mathbf{r}') \exp\left(-ik\hat{\mathbf{r}} \cdot \mathbf{r}'\right) \, \mathrm{d}\mathcal{S}' \quad kr \to \infty. \tag{9.49}$$

The next major simplification results from approximating the induced surface current $\mathbf{K}_b$. One possibility would be to assume that this current is the same as that which would occupy the same area as $\mathcal{S}_a^{(b)}$ on the complete infinite plate. This is equivalent to replacing the total field at the surface $\mathcal{S}_a^{(b)}$ in configuration (b) by the incident field only, and is the essence of the Kirchhoff approximation. Notice the similarity with this approximation and the Rayleigh-Debye-Gans approximation of Ch. 7. Then, with the help of Eq. (9.7)

$$\mathbf{K}_b(\mathbf{r}) = -\frac{2k}{\omega\mu_o} \mathbf{E}_o^{inc} \quad \mathbf{r} \in \mathcal{S}_a^{(b)},$$

or, more explicitly,

$$\mathbf{K}_b(\mathbf{r}) = -\frac{2k}{\omega\mu_o} E_o^{inc} \, \hat{\mathbf{y}} \quad \mathbf{r} \in \mathcal{S}_a^{(b)}. \tag{9.50}$$

Equations (9.48) and (9.49) then

$$\mathbf{E}_a(\mathbf{r}) \simeq 2ik\varepsilon_o E_o^{inc} \frac{\exp(ikr)}{4\pi r} \left(\hat{\mathbf{r}} \times \hat{\mathbf{y}}\right) \int_{\mathcal{S}_a^{(b)}} \exp\left(-ik\hat{\mathbf{r}} \cdot \mathbf{r}'\right) \, \mathrm{d}\mathcal{S}', \tag{9.51}$$

and

$$\mathbf{B}_a(\mathbf{r}) \simeq -2i\frac{k}{\omega} E_o^{inc} \frac{\exp(ikr)}{4\pi r} \left[\hat{\mathbf{y}} - \left(\hat{\mathbf{r}} \cdot \hat{\mathbf{y}}\right) \hat{\mathbf{r}}\right] \int_{\mathcal{S}_a^{(b)}} \exp\left(-ik\hat{\mathbf{r}} \cdot \mathbf{r}'\right) \, \mathrm{d}\mathcal{S}'. \tag{9.52}$$

A check of these expressions is to expand the aperture-obstacle $\mathcal{S}_a^{(b)}$ until it becomes an infinite plate, which corresponds to the screen vanishing in configuration (a). One should then recover the incident field from Eqs. (9.51) and (9.52). Notice, thought, that expanding $\mathcal{S}_a^{(b)}$ in this way violates the far-field conditions of Eq. (4.13)-(4.15), which are used to derive Eqs. (9.51) and (9.52). With the help of Sec. 4.2, it can be shown that the term

$$\frac{\exp(ikr)}{r} \int_{\mathcal{S}_a^{(b)}} \exp\left(-ik\hat{\mathbf{r}} \cdot \mathbf{r}'\right) \, \mathrm{d}\mathcal{S}',$$

appearing in Eqs. (9.51) and (9.52) is the far-field approximation of

$$\int_{\mathcal{S}_a^{(b)}} \frac{\exp(ik\mathfrak{r})}{\mathfrak{r}} \, \mathrm{d}\mathcal{S}'$$

where $\mathfrak{r} = \mathbf{r} - \mathbf{r}'$. Also remember Fig. (2.1). Then,

$$\int_{\mathcal{S}_a^{(b)} \to \infty} \frac{\exp(ik\mathfrak{r})}{\mathfrak{r}} \, \mathrm{d}\mathcal{S}' = -\frac{2\pi}{ik} \exp(ikz), \tag{9.53}$$

where use has been made of the specific orientation of the integration surface, i.e. $\mathcal{S}_a^{(b)}$ coinciding with the x-y plane through the origin. With this consideration and the calculations in Sec. 2.3.2, one can eventually show that Eqs. (9.51) and (9.52) become

$$\mathbf{E}_a(\mathbf{r}) = E_o^{inc} \exp(ikz) \, \hat{\mathbf{x}},$$

and

$$\mathbf{B}_a(\mathbf{r}) = \frac{k}{\omega} E_o^{inc} \exp(ikz) \, \hat{\mathbf{y}},$$

which are the incident fields as expected.

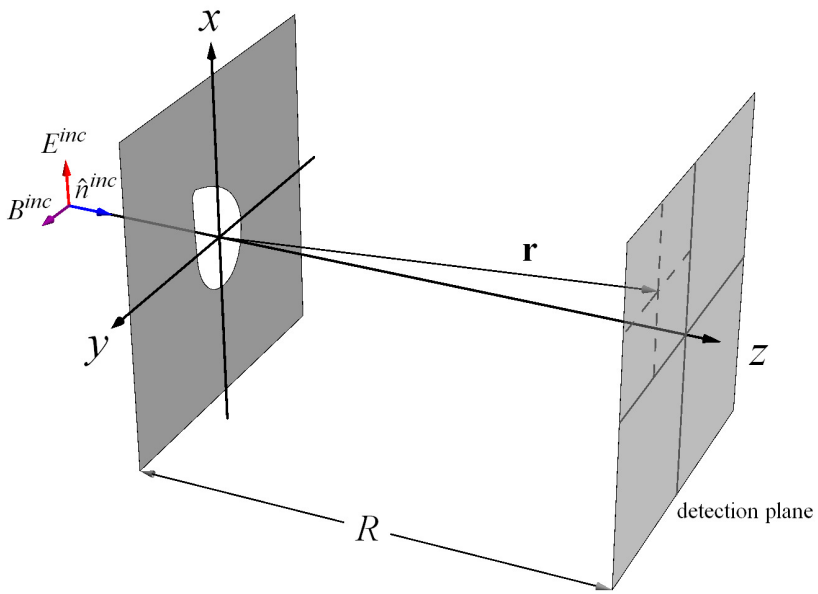

Figure 9.6: Detector plane.

Now consider the energy flow at the observation point due to the scattered wave, given by Eq. (3.28)

$$\langle \mathbf{S}^{sca}(\mathbf{r}) \rangle_t = \frac{1}{2\mu_o} \mathrm{Re}\Big\{ \mathbf{E}^{sca}(\mathbf{r}) \times [\mathbf{B}^{sca}(\mathbf{r})]^* \Big\}.$$

Combining this with Eqs. (9.51) and (9.52) gives

$$\langle \mathbf{S}^{sca}(\mathbf{r})\rangle_t = \frac{1}{2\pi}\left(\frac{E_o^{inc}}{r}\right)^2 \sqrt{\frac{\varepsilon_o}{\mu_o}}\,|F(\hat{\mathbf{r}})|^2\left[\hat{\mathbf{r}} - (\hat{\mathbf{r}}\cdot\hat{\mathbf{y}})^2\,\hat{\mathbf{r}}\right], \tag{9.54}$$

where

$$F(\hat{\mathbf{r}}) = \int_{\mathcal{S}_a^{(b)}} \exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathcal{S}'. \tag{9.55}$$

Equation (9.54) can be cast into a far more familiar form by considering only the component of the energy flow that is directed into a detection plane parallel to the screen and separated from it by a distance R as shown in Fig. (9.6). This component is the z-component, and so, the intensity of the far-field scattered wave on the detector plane is given by

$$I(\mathbf{r}) = R^2\langle \mathbf{S}^{sca}(\mathbf{r})\rangle_t \cdot \hat{\mathbf{z}}, \tag{9.56}$$

which when combined with Eqs. (9.54) and (9.55) gives

$$I(\mathbf{r}) = \frac{1}{2\pi}\left(E_o^{inc}\right)^2 \sqrt{\frac{\varepsilon_o}{\mu_o}}\,|F(\hat{\mathbf{r}})|^2\,. \tag{9.57}$$

Now notice that since the screen is in the $z = z' = 0$ plane, Eq. (9.55) can be expressed as

$$F(\hat{\mathbf{r}}) = \int_{\mathcal{S}_a^{(b)}} \exp(ikz' - ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathcal{S}'. \tag{9.58}$$

But

$$z' = \mathbf{r}'\cdot\hat{\mathbf{n}}_a^{inc},$$

which when combined with Eq. (9.58) gives

$$F(\hat{\mathbf{r}}) = \int_{\mathcal{S}_a^{(b)}} \exp\left[ik\mathbf{q}(\hat{\mathbf{r}})\cdot\mathbf{r}'\right]\,\mathrm{d}\mathcal{S}', \tag{9.59}$$

where $\mathbf{q}$ is the scattering wave vector of Eq. (7.2). Now recall the Fourier transform pair of Eqs. (1.16) and (1.17) in the context of Eq. (9.59),

$$F(\mathbf{q}) = \frac{1}{2\pi}\int_{\mathbb{R}^2} f(\mathbf{r}')\exp\left[ik\mathbf{q}(\hat{\mathbf{r}})\cdot\mathbf{r}'\right]\,\mathrm{d}\mathbf{r}', \tag{9.60}$$

and

$$f(\mathbf{r}) = \frac{1}{2\pi}\int_{\mathbb{Q}^2} F(\mathbf{q})\exp\left[-ik\mathbf{q}(\hat{\mathbf{r}})\cdot\mathbf{r}\right]\,\mathrm{d}\mathbf{q}, \tag{9.61}$$

where $\mathbb{R}^2$ and $\mathbb{Q}^2$ denote all of the two-dimensional real and q-space. Next define an aperture function A as

$$A(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in \mathcal{S}_a^{(b)} \\ 0 & \mathbf{r} \notin \mathcal{S}_a^{(b)}. \end{cases} \tag{9.62}$$

Then Eq. (9.59) can be expressed as

$$F(\mathbf{q}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} A(\mathbf{r}') \exp\left[ik\mathbf{q}(\hat{\mathbf{r}}) \cdot \mathbf{r}'\right] \, \mathrm{d}\mathbf{r}', \tag{9.63}$$

so

$$\frac{I(\hat{\mathbf{r}})}{I(\hat{\mathbf{n}}^{inc})} = \left| \int_{\mathbb{R}^2} A(\mathbf{r}') \exp\left[ik\mathbf{q}(\hat{\mathbf{r}}) \cdot \mathbf{r}'\right] \, \mathrm{d}\mathbf{r}' \right|^2, \tag{9.64}$$

which is *Huygens principle*. Notice that these equations describe the far-field diffracted intensity in a way that appears as though there are wavelet sources covering the aperture opening; these are the fictitious sources discussed at the beginning of this chapter.

Equation (9.64) is the classic result that the far-field distribution of intensity due to the aperture is given by the aperture's Fourier transform into q-space. The major elements ultimately responsible for this result are Babinet's principle and the far-field limit. Babinet's principle permits a *mathematically* equivalent description of the fields radiated by the screen's current as the field due to the currents in the complimentary problem. This equivalence is due to the duality of the Maxwell equations *and* the infinite-in-extent two-dimensional planar geometry of the screen. The fictitious spherical wave sources that span the aperture in the Huygens description are a consequence of this mathematical connection between the *real* currents induced on the screen and its fictitious compliment. The appearance of the Fourier transform in Huygens' principle is a consequence of the VIE (9.11) and the form that the Green's function takes in the far-field limit.

9.5 Examples

As mentioned above, the DDA can be used to find the current induced on the screen and calculate the scattered wave. Because of the infinite extent of the screen this must be done indirectly via Babinet's principle by applying the DDA to the complimentary configuration

(b) in Fig. (9.5). Application of the DDA to these complimentary configurations is done by discretizing the aperture-obstacles $\mathcal{S}_a^{(b)}$ and $\mathcal{S}_a^{(d)}$ on a square grid into square surface elements $\Delta\mathcal{S}_i$ represented by electric dipole moments $\mathbf{p}_i$, recall Ch. 5. The polarizability of the dipole moments is given by Eq. (5.12) in the limit $m \to \infty$, i.e., the PC limit. Once the coupling between the dipoles has been determined by solving the DDA matrix equation (5.15), the dipole moments are related to the induced surface current in analogy to Eq. (2.53) as

$$\mathbf{K}(\mathbf{r}_i)\Delta\mathcal{S}_i = -i\omega\mathbf{p}_i. \tag{9.65}$$

Figure (9.7) shows an infinite screen with a powercat aperture. The illumination configuration is the same as Fig. (9.3) and the detector plane is the same as Fig. (9.6). The largest dimension of the aperture is $l = 8.6\lambda$. Shown in Fig. (9.8) is the deviation of the induced surface current *on the aperture-obstacle*, $\mathbf{K}^{(a)}$, from the Kirchhoff expectation. The color code in the plot displays the magnitude of this deviation,

$$\frac{|\mathbf{K}^{inc}(\mathbf{r})| - |\mathbf{K}^{(b)}(\mathbf{r})|}{|\mathbf{K}^{inc}(\mathbf{r})|}, \tag{9.66}$$

where $\mathbf{K}^{inc}$ is the current that would be induced by the incident field only. The vectors shown in Fig. (9.8) display the direction of the full induced current, i.e., not the direction of the Kirchhoff-relative current deviation. Notice that the current-deviation on the complimentary aperture-obstacle points along the y-axis. This is because of the rotation of the polarization of the incident wave in the complimentary configurations in Babinet's principle. One can see that, overall, the current deviations are greatest near regions with sharp points.

The current on the Powercat *screen*, $\mathbf{K}^{(a)}$, is found from the DDA via Eqs. (9.42) and (9.65). In analogy to Fig. (9.8), the Kirchhoff-relative current deviation for the screen current is shown in Fig. (9.9). The color shades in this case are calculated from Eq. (9.66) with $\mathbf{K}^{(a)}$ substituted for $\mathbf{K}^{(b)}$ in that equation. Remember that the vectors in this plot show the direction of the current, not of the deviation. The qualitative behavior of the current here is similar to that of the aperture-obstacle, except one can also see a rippled

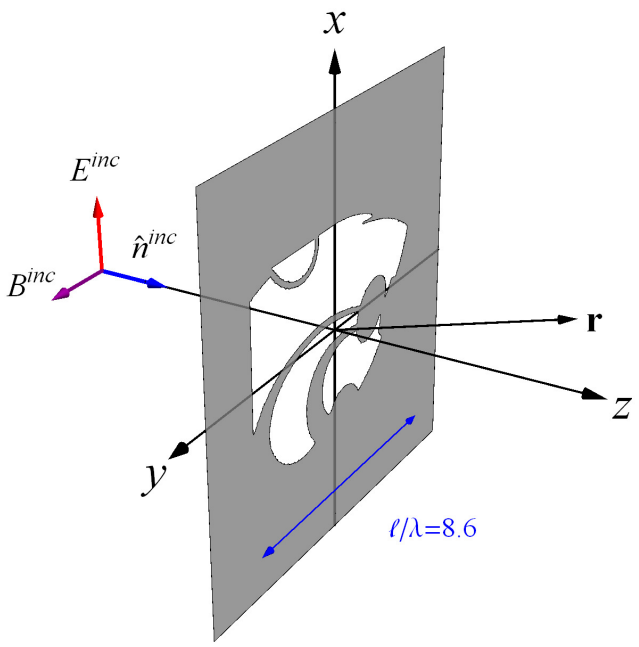

Figure 9.7: Powercat shaped aperture in an infinite perfectly conducting screen.

structure emanating away from the aperture-opening with spacing that is roughly the same as the vacuum wavelength.

Figures (9.10) and (9.11) show the near and far-field scattered wave intensity on the detector plane shown in Fig. (9.6). In the near-field, the screen-to-detector separation is $R = \lambda$, and in the far-field case the separation is $R = 425\lambda$. One can verify that this latter case satisfies Eqs. (4.13)-(4.15). Superimposed on the intensity distributions is the outline of the projection of the aperture onto the detector plane. In the far-field case, Fig. (9.11), one can see that the expected qualitative Fourier transform like behavior of the intensity distribution is present; it is narrowed along the widest aspect of the aperture and widened along the narrowest aspect of the aperture.

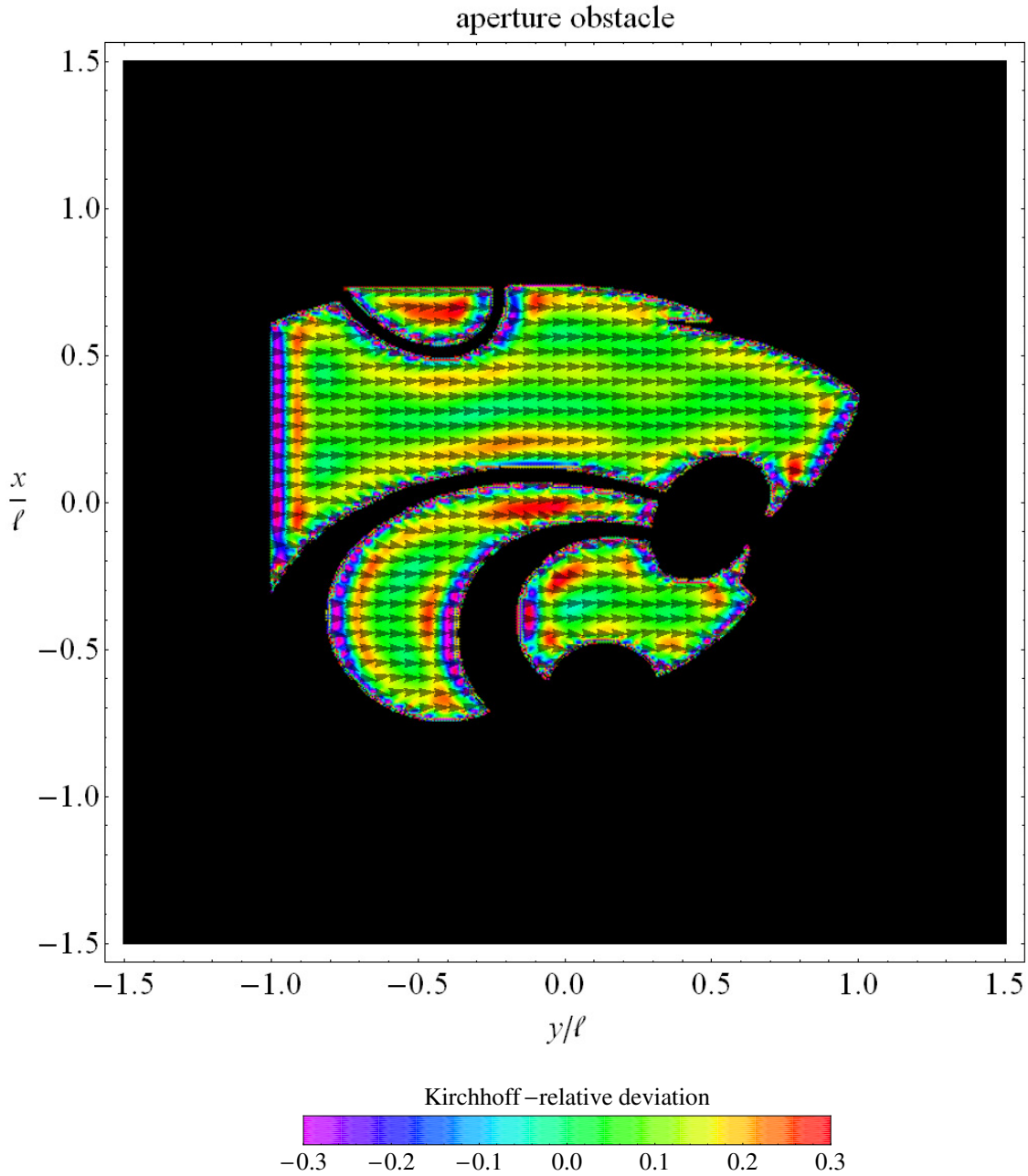

Figure 9.8: Kirchhoff-relative surface current induced on the complimentary Powercat aperture-obstacle corresponding to Fig. (9.7). The color shades indicate the current magnitude deviation according to Eq. (9.66), whereas the vectors indicate the direction of the current, not the direction of the deviation.

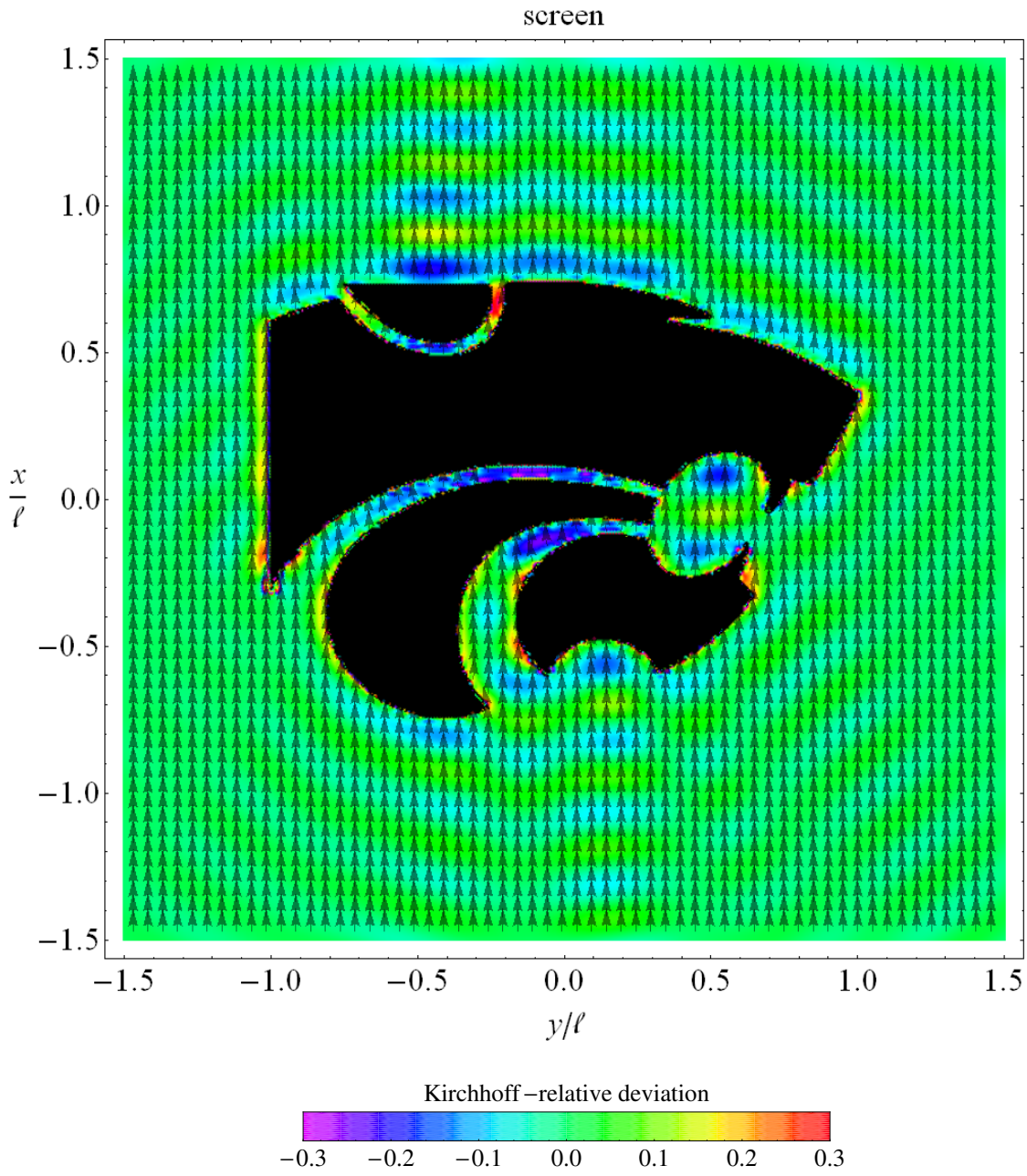

Figure 9.9: Kirchhoff-relative surface current induced on the Powercat screen of Fig. (9.7). Recall the discussion in Fig. (9.8) relating to Eq. (9.66).

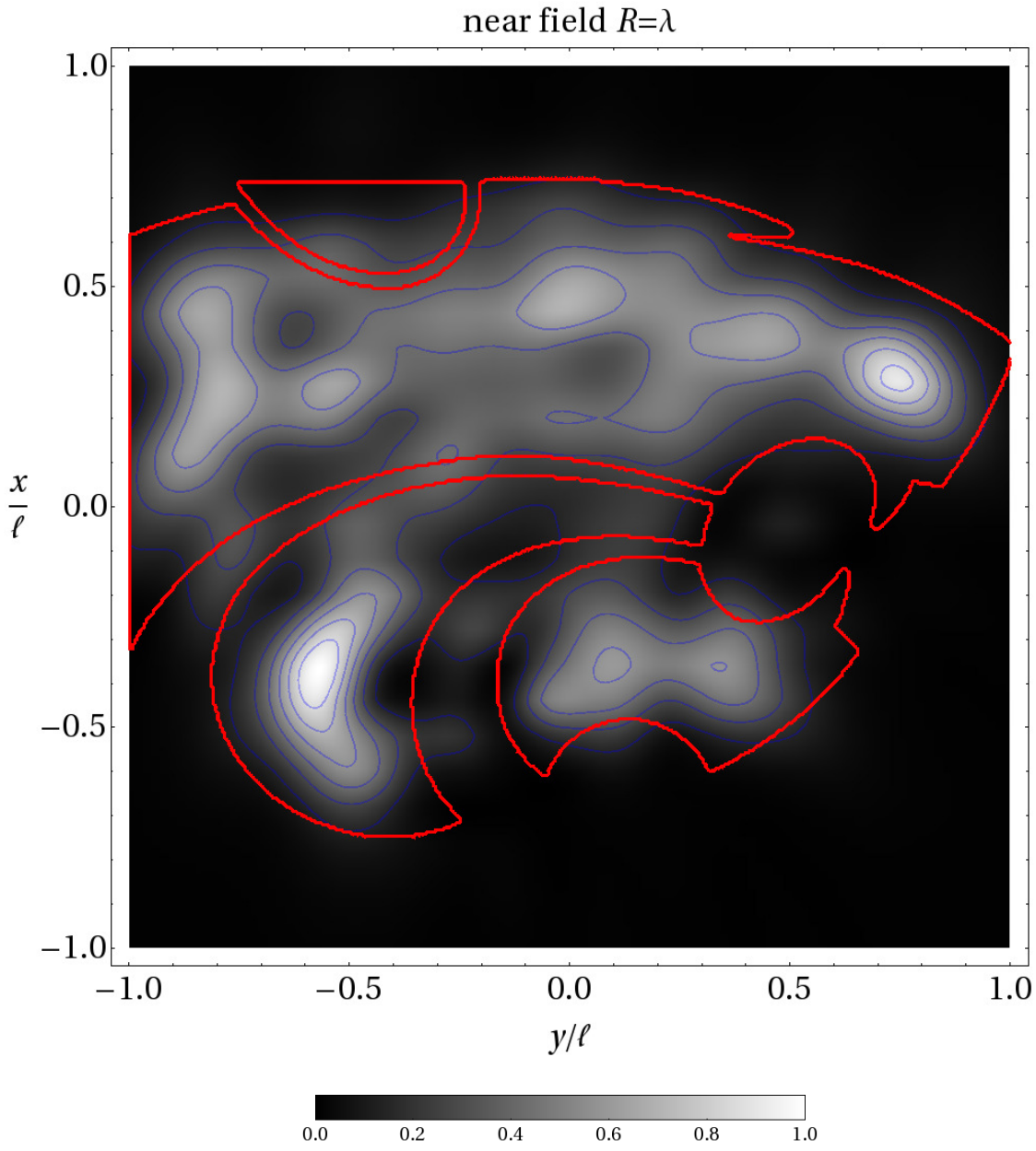

Figure 9.10: Near-field diffraction from Powercat aperture. The detector plane is separated from the screen by $R = \lambda$ as shown in Fig. (9.6). The red outline represents the projection of the Powercat aperture onto the detector screen, recall Fig. (9.7).

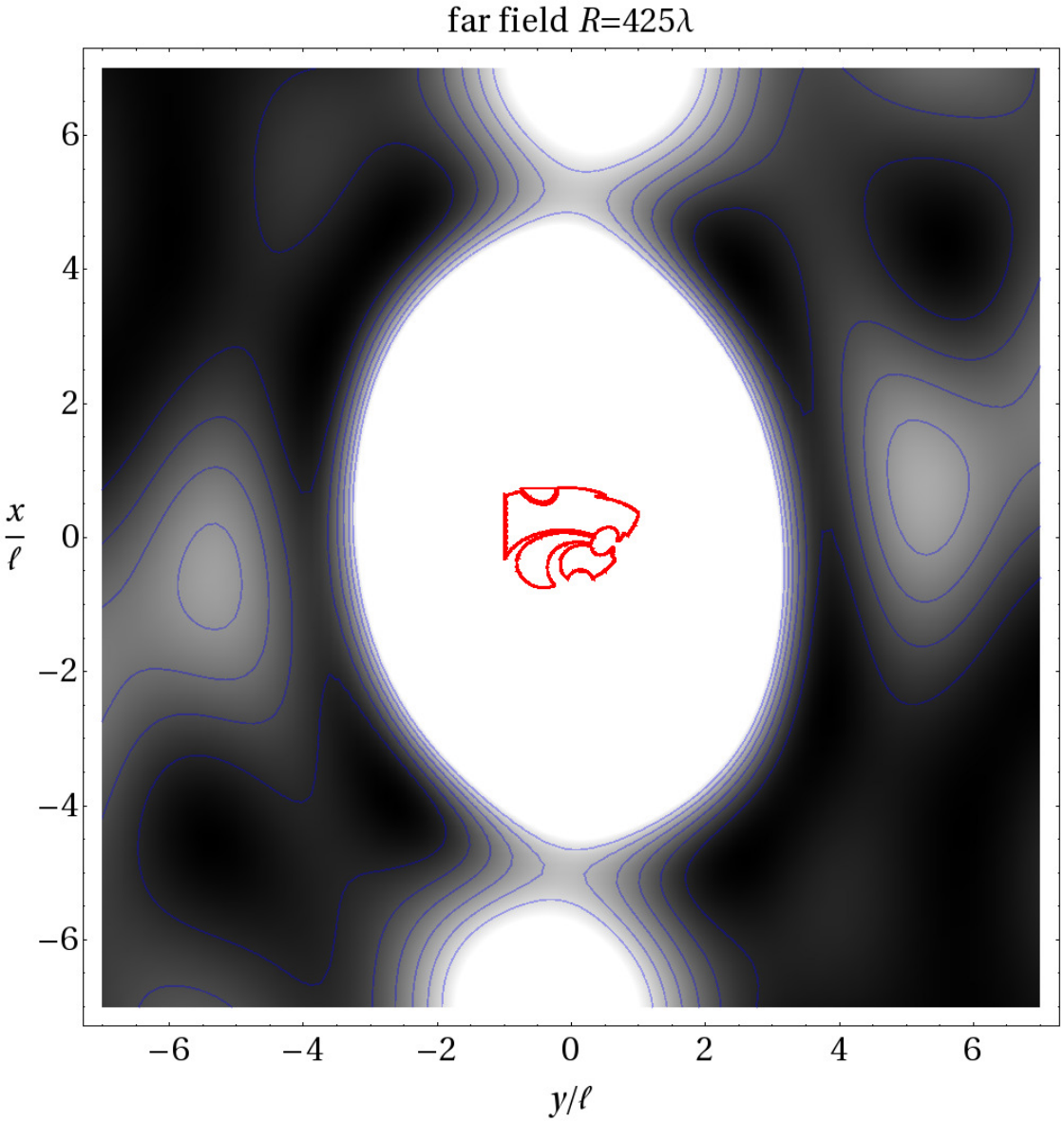

Figure 9.11: Same as in Fig. (9.10) except for the far-field zone, where the screen-detector separation is $R = 425\lambda$.

9.6 Discussion

Diffraction is typically described as the deviation of light from rectilinear propagation as it passes by a edge of an opaque obstacle [59, p. 443]. In other words, diffraction is the bending of light around an edge or corner. Unlike the bending associated with refraction, diffraction usually refers to this bending as it occurs in vacuum. This definition of diffraction is basically phenomenological. The calculations and concepts in this chapter, however, provide one with a more physically explicit explanation for how the physical effects associated with diffraction occur.

Recall Sec. 9.1 where the fields on either side of an infinite perfectly conducting thin plate are found. The total fields behind the plate are zero, not because the plate "blocks" the wave, but because the wave induces a surface current on the plate that is just right in direction and magnitude such that it radiates a secondary wave that exactly cancels the incident wave. The energy density eliminated from the half-space behind the plate $V^{(+)}$ due to this cancellation appears on the front side $V^{(-)}$ in the form of a standing wave with twice the magnitude (and energy density) as the incident wave. Thus, energy is conserved, and the redistribution of the energy is facilitated by the phase shift in the induced current relative to the incident wave.

When an aperture is cut into the infinite plate, making it the screen, the current induced on the the screen is disturbed from what is otherwise expected. This is what is seen in Fig. (9.9). The absence of current in the aperture means that the delicate cancellation of the incident wave becomes incomplete resulting in a nonzero wave magnitude behind the screen. This is the unorthodox view of diffraction. The "bending" of light around the aperture edge is actually a consequence of the superposition of the incident wave with the wavelets radiated from the differential surface elements of induced current everywhere on the screen.

Scattering and diffraction are often referred to as though they represent different physical effects: Sunlight *scatters* from an ice crystal in the atmosphere, but a laser beam *diffracts* from a thin slit in a metal sheet. The microphysical view of diffraction makes no such

distinction. The reality is that scattering and diffraction are not fundamentally separate physical effects but are one in the same; both are consequences of the way that the object's wavelet sources couple and radiate secondary waves that superimpose to establish the observed fields. Although this view of diffraction and scattering may seem somewhat pedantic, one will see in Ch. 11 that retaining the mindset that diffraction and scattering are somehow separate effects prevents one from correctly explaining the cause of the extinction paradox.

Chapter 10

Extinction and the Optical Theorem

"...the actual extinction process is not a blocking of the wave but a subtle interference phenomenon."

-H. C. van de Hulst [47]

Chapter 3 introduces and derives general expressions for the total cross sections for an arbitrary particle. The considerations there are largely mathematical in nature and touch on only several aspects of the physical significance of the cross sections. The work in this chapter revisits the cross sections, especially the extinction cross section, but does so in a context more aligned with measurements. The extinction due to a single particle is presented, and a conceptual phase-based explanation for the related optical theorem is developed. Simulations of the energy flow caused by a particle in a beam of light demonstrate how the extinction process occurs. It is shown that extinction does not necessarily cause a reduction of the energy flow along the exact forward direction, which is contrary to how extinction is typically described. Implications regarding the measurement of the cross sections are discussed. The single-particle work is extended to non-interacting and fully-interacting multi-particle groups. From this one sees how the forward angle character of the well-known Beer's law is a consequence of the presence of multiple particles.

10.1 Theory and Meaning of Extinction

Consider the situation depicted in Fig. (10.1). The sketch labeled (a) shows a detector with its sensitive face exposed to the energy flow of an oncoming collimated beam of light. This beam is considered the incident wave and propagates in the "forward direction." Now consider the sketch labeled (b). Here a particle is present as indicated by P. The detector is assumed to be in the particle's far-field zone. With the particle present, the incident wave no longer exists due interactions with the particle as described in Sec. 1.7; the wave is replaced by a new wave, referred to as the total wave, that propagates and carries energy in all directions and hence accounts for the scattering of the incident beam. Energy can also be absorbed by the particle's material and this absorption accounts for a net flow of energy into the particle's interior. Energy flows due to scattering and absorption are both figuratively illustrated in Fig. 10.1(b) and are associated with a reduction of the power received by the detector facing into the forward direction as shown. *Extinction is the effect that is associated with the reduction of the power received by this detector when a particle is present.*

The essence of extinction is the conservation of energy. Analogs of the above example and the related optical theorem occur in acoustical and quantum-mechanical scattering [92]. Extinction has received extensive consideration in the literature because of its fundamental nature [2, p. 3, Secs. 2.5, 2.8], [47, Secs. 4.21, 4.22, 4.3], [37, Secs. 3.4, 11], [42], [93, 94], [95, Ch. 5]. Despite this attention, however, a careful graphically-based examination of how extinction occurs and the subtle details of its theoretical description, including the optical theorem, have yet to be presented. The exposition given in this chapter will study extinction in the context of single and multi-particle systems and will result in a new understanding. For example, extinction is often described as the reduction of the energy flow along the forward direction due to scattering and absorption. Indeed, this is consistent with the operational definition of extinction embodied by Fig. (10.1) above. However, it will be shown that extinction *does not* necessarily result in a reduction of the total energy flow

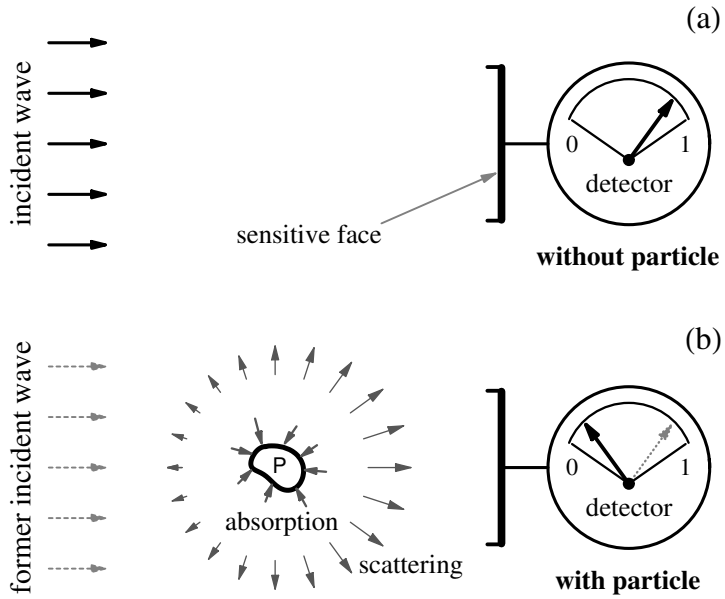

Figure 10.1: Effect of extinction as described by two measurements. Sketch (a) shows a detector with its sensitive face looking into the incident beam when no particle is present. Sketch (b) shows the same situation with a particle present and labeled by P. The energy flows corresponding to the incident and scattered waves, and the flow into the particle due to absorption, are both figuratively indicated.

along the *exact* forward direction, and can sometimes result in the enhancement of this flow. This forward-flow behavior is not only consistent with, but is requisite for, the conservation of energy. The measurements pertaining to Fig. (10.1) will be revisited, and the familiar characteristic of extinction as reducing the power received by a detector will be shown to be a consequence of the detector having a finite size.

10.2 Mathematical Background

Suppose that a single particle, which is illuminated by the incident wave of Eqs. (3.1) and (3.2), is centered at the origin and surrounded by vacuum[1]. Surrounding the particle is the large imaginary spherical surface $\mathcal{S}_{en}$ of radius R_{en} introduced in Sec. 3.1. Here, however, this surface is large enough that points on it satisfy the far-field conditions of Eqs. (4.13)-(4.15) and $\mathbf{r}$ is restricted to points on $\mathcal{S}_{en}$ in the following for simplicity.

[1]See [38, 41, 96] for generalizations to the case when the particle resides in an absorbent medium.

Because $\mathbf{r}$ satisfies the far-field conditions on $\mathcal{S}_{en}$, the scattered wave is a transverse outward-traveling spherical wave, recall Sec. 2.3.1. This results in the simple scattering amplitude form for the scattered electric field given by Eq. (4.17)

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{\exp(ikr)}{r}\mathbf{E}_o^{sca}(\hat{\mathbf{r}}). \tag{10.1}$$

Note that the scattering amplitude $\mathbf{E}_o^{sca}$ is exactly the same as $\mathbf{E}_1^{sca}$ of Eq. (4.17); the subscript has been changed to avoid notational confusion when dealing with multiple-particle groups later. The magnetic field follows from Eq. (10.1) via the transverse nature of the far-field scattered wave,

$$\mathbf{B}^{sca}(\mathbf{r}) = \frac{k}{\omega}\frac{\exp(ikr)}{r}\,\hat{\mathbf{r}} \times \mathbf{E}_o^{sca}(\hat{\mathbf{r}})\,. \tag{10.2}$$

Now the microphysical model is introduced by expressing the scattering amplitude in terms of the far-field limit of the VIE, i.e. Eq. (7.6),

$$\mathbf{E}^{sca}(\mathbf{r}) = \frac{k^2}{4\pi}\frac{\exp(ikr)}{r}(m^2-1)\int_{V^{int}}\left(\overleftrightarrow{\mathbf{I}} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\mathbf{E}^{int}(\mathbf{r}')\exp(-ik\hat{\mathbf{r}}\cdot\mathbf{r}')\,\mathrm{d}\mathbf{r}' \quad kr\to\infty,$$

which from Eq. (10.1) gives

$$\mathbf{E}_o^{sca}(\hat{\mathbf{r}}) = \frac{k^2}{4\pi}\left(m^2-1\right)\left(\overleftrightarrow{I} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\cdot\int_V \mathbf{E}^{int}(\mathbf{r}')\,\exp(-ik\,\mathbf{r}'\cdot\hat{\mathbf{r}})\,\mathrm{d}\mathbf{r}'. \tag{10.3}$$

Recall from Sec. 4.2, that Eqs. (10.1)-(10.3) effectively describe the particle as a point at the origin, where the entire angular variability of the fields is now contained in the r-independent scattering amplitude $\mathbf{E}_o^{sca}$.

The transport of energy in an electromagnetic wave is given by the Poynting vector, which in its time-averaged form is related to the fields by Eq. (1.45). The units of the Poynting vector are energy/(area×time) and hence describe an energy flux, or flow. In the following when comparisons to measurement are made, the component of $\langle\mathbf{S}\rangle_t$ directed into the sensitive face of a detector and integrated over the face will be taken to represent the detector's response, recall Sec. 1.5.

From Eqs. (3.3) and (3.4), Eq. (1.45) factors into three distinct terms

$$\langle\mathbf{S}(\mathbf{r})\rangle_t = \langle\mathbf{S}^{inc}(\mathbf{r})\rangle_t + \langle\mathbf{S}^{sca}(\mathbf{r})\rangle_t + \langle\mathbf{S}^{cross}(\mathbf{r})\rangle_t, \tag{10.4}$$

where,

$$\langle \mathbf{S}^{inc}(\mathbf{r})\rangle_t = \frac{1}{2\mu_o}\,\mathrm{Re}\Big\{\mathbf{E}^{inc}(\mathbf{r}) \times \big[\mathbf{B}^{inc}(\mathbf{r})\big]^* \Big\}, \tag{10.5}$$

$$\langle \mathbf{S}^{sca}(\mathbf{r})\rangle_t = \frac{1}{2\mu_o}\,\mathrm{Re}\Big\{\mathbf{E}^{sca}(\mathbf{r}) \times [\mathbf{B}^{sca}(\mathbf{r})]^* \Big\}, \tag{10.6}$$

$$\langle \mathbf{S}^{cross}(\mathbf{r})\rangle_t = \frac{1}{2\mu_o}\,\mathrm{Re}\Big\{\mathbf{E}^{inc}(\mathbf{r}) \times [\mathbf{B}^{sca}(\mathbf{r})]^* + \mathbf{E}^{sca}(\mathbf{r}) \times \big[\mathbf{B}^{inc}(\mathbf{r})\big]^* \Big\}. \tag{10.7}$$

Equation (10.5) describes the energy flow at $\mathbf{r}$ due to the incident wave and Eq. (3.28), which is identical to Eq. (3.28), describes the flow due to the scattered wave. Notice that the cross term of Eq. (10.7) involves products of the fields of both the incident and scattered waves; this shows that $\langle \mathbf{S}^{cross}\rangle_t$ describes the portion of the total energy flow at $\mathbf{r}$ due to the *interference* of the two waves.

Electromagnetic energy will be removed from the wave if the particle's material is absorbent. This loss constitutes a net inward flow through $\mathcal{S}_{en}$ given by [2, p. 57]

$$W^{abs} = -\oint_{\mathcal{S}_{en}} \langle \mathbf{S}(\mathbf{r})\rangle_t \cdot \hat{\mathbf{r}}\, \mathrm{d}S\,. \tag{10.8}$$

The quantity W^{abs} gives the total power absorbed within the particle and is either positive or zero. Using Eq. (10.4), this becomes

$$W^{abs} = W^{inc} - W^{sca} + W^{cross}, \tag{10.9}$$

where

$$W^{inc} = -\oint_{\mathcal{S}_{en}} \langle \mathbf{S}^{inc}(\mathbf{r})\rangle_t \cdot \hat{\mathbf{r}}\, \mathrm{d}S\,, \tag{10.10}$$

$$W^{sca} = \oint_{\mathcal{S}_{en}} \langle \mathbf{S}^{sca}(\mathbf{r})\rangle_t \cdot \hat{\mathbf{r}}\, \mathrm{d}S\,, \tag{10.11}$$

$$W^{cross} = -\oint_{\mathcal{S}_{en}} \langle \mathbf{S}^{cross}(\mathbf{r})\rangle_t \cdot \hat{\mathbf{r}}\, \mathrm{d}S\,. \tag{10.12}$$

The quantity W^{inc} gives the net power crossing $\mathcal{S}_{en}$ due to the incident wave, which is zero since the incident wave carries as much energy per unit time into the volume bounded by $\mathcal{S}_{en}$ as it carries out of the volume, see [44, p. 50] and [2, p. 57]. The quantity W^{sca} gives

the net power crossing $\mathcal{S}_{en}$ due to the outward-traveling scattered wave and is either positive or zero. With $W^{inc} = 0$, Eq. (10.9) now reads

$$W^{cross} = W^{sca} + W^{abs}. \tag{10.13}$$

This is the *conservation of energy* as it pertains to the energy content of the scattered wave, the interference between the incident and scattered waves, and the energy absorbed by the particle.

To understand the meaning of W^{cross} let the volume of space inside of $\mathcal{S}_{en}$, including the particle, define the so-called system [6, p. 77-78]. Equation (10.13) then shows that W^{cross} is equal to the net power lost by the system due the outward traveling scattered wave and absorption within the particle. *This net loss is extinction and it is a statement of the conservation of energy.* Notice that the definition of extinction makes *no explicit reference to any requisite attenuation of the net energy flow along the forward direction.* This is evident from the clear dependence in Eqs. (10.11) and (10.12) on *all* directions.

The scattering, absorption, and extinction cross sections C^{sca}, C^{abs} and C^{ext} are defined from W^{sca}, W^{abs}, and W^{cross}, respectively, by normalizing the latter by the energy flux of the incident wave $I^{inc} = (\epsilon_o/\mu_o)^{1/2}|\mathbf{E}_o^{inc}|^2/2$. An important relation for the extinction cross section, the optical theorem, is derived by combining Eqs. (10.7) and (10.12) to give

$$W^{cross} = -\frac{1}{2\mu_o}\,\mathrm{Re}\oint_{\mathcal{S}_{en}} \left\{\mathbf{E}^{inc}(\mathbf{r}) \times [\mathbf{B}^{sca}(\mathbf{r})]^* + \mathbf{E}^{sca}(\mathbf{r}) \times \left[\mathbf{B}^{inc}(\mathbf{r})\right]^* \right\} \cdot \hat{\mathbf{r}}\,\mathrm{d}S\,. \tag{10.14}$$

The plane wave exponentials that appear in Eq. (10.14) through $\mathbf{E}^{inc}$ and $\mathbf{B}^{inc}$, recall Eqs. (3.1) and (3.2), can be expanded as

$$\exp(ikr\hat{\mathbf{r}} \cdot \hat{\mathbf{n}}^{inc}) = 4\pi \sum_{l=0}^{\infty} i^l\, j_l(kr) \sum_{m=-l}^{l} Y_{lm}^*(\hat{\mathbf{r}}) Y_{lm}(\hat{\mathbf{n}}^{inc}), \tag{10.15}$$

where j_l are spherical Bessel functions of the first kind and order l, and Y_{lm} are spherical harmonics, see [2, App. A]. By taking the limit that the observation point goes to infinity, the large-argument form of the spherical Bessel functions can be used in combination with

the completeness relation for the spherical harmonics to reduce Eq. (10.15) to

$$\exp(ikr\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}^{inc}) = 2\pi i\Big[\delta(\hat{\mathbf{r}}+\hat{\mathbf{n}}^{inc})\,\frac{exp(-ikr)}{kr} - \delta(\hat{\mathbf{r}}-\hat{\mathbf{n}}^{inc})\,\frac{exp(ikr)}{kr}\Big] \quad kr\to\infty. \tag{10.16}$$

Combining Eq. (10.14) with Eq. (10.16) using Eqs. (3.1), (3.2), (10.1), and (10.2), and remembering that $C^{ext} = W^{cross}/I^{inc}$ shows that the extinction cross section can be expressed as

$$C^{ext} = \frac{4\pi}{k|\mathbf{E}_o^{inc}|^2}\,\mathrm{Im}\big\{\mathbf{E}_o^{inc*}\cdot\mathbf{E}_o^{sca}(\hat{\mathbf{n}}^{inc})\big\}. \tag{10.17}$$

Equation (10.17) is the optical theorem; it relates the extinction cross section to the imaginary part of the scattering amplitude evaluated in the forward direction. An alternative way to arrive at the optical theorem from Eq. (10.14) is to use Jones' lemma, which utilizes the method of stationary phase to evaluate Eq. (10.14) in the limit that $kr\to\infty$, see [24, App. 12]. Regardless of how it is derived, the optical theorem requires the limit $kr\to\infty$. This point serves as a reminder that *the theorem is only approximate in the far-field and becomes strictly valid only at infinity*. In addition, the incident wave must be planar otherwise Eq. (10.17) can fail, e.g., see [97].

10.3 Subtle Nature of Extinction

Much of the present understanding of extinction and the physical meaning of the optical theorem traces to van de Hulst [47, 98]. In that work, the removal of energy from the incident wave is claimed to be due to the interference of the incident and scattered waves in *only* the "neighborhood of the forward direction." When looking closer however, several questions arise. If extinction is to be regarded as the combined effect of scattering and absorption, as Eq. (10.13) implies, then

- why does the optical theorem show no *apparent* account for absorption, and
- why does the optical theorem involve only the forward direction when the scattered wave, upon which the theorem depends, has energy flow in all directions?

Here the answers to these questions are summarized and the sections to follow give detail to these answers.

The dependence of extinction on a particle's absorption is hidden in the scattering amplitude $\mathbf{E}_o^{sca}$. The VIE shows that the scattering amplitude depends on the specific structure of the particle's internal electric field. The magnitude and direction of this field is affected by the absorptive properties of the particle's material.

The forward direction dependence of the optical theorem is a result of the way that the incident and scattered waves interfere. Equations (10.7), (10.12) and (10.13) show that the extinction caused by the particle is given by the component of the interference energy flow that passes through $\mathcal{S}_{en}$ in all directions. Below it is shown that this flow alternates between inward and outward through $\mathcal{S}_{en}$ as a function of direction. Moreover, it is shown that the angular distribution of the alternating flow changes with distance from the particle in the far-field in every direction except the exact forward direction. Because a particle's extinction cross section must ultimately be independent of the distance from the particle, as proven in Ch. 3, one can conclude that the optical theorem should then depend only on the forward direction, cf. [99, 100].

10.4 Interference

Consider the formulation of Eq. (10.12) as being the cross term between the incident and scattered waves, i.e., the interference of these two waves [95]. Figure (10.2) qualitatively shows how the planar incident wave and the scattered spherical wave interfere[2] in the far-

[2]Confusion can arise when trying to identify the surfaces of constant phase of the far-field scattered wave. Clearly, the term $\exp(ikr)/r$ in Eq. (10.1) is constant in both magnitude and phase on spherical surfaces centered on the origin. One can see from Eq. (10.3) that $\mathbf{E}_o^{sca}$ will, in general, vary in magnitude and phase on these same spherical surfaces. One could take the surfaces of constant phase of the wave to be spheres and then regard $\mathbf{E}_o^{sca}$ as a complex-valued angular weighting function that varies over these surfaces. Alternatively, one could use Euler's formula to express the right-hand side of Eqs. (10.1) or (10.2) as $[A(\hat{\mathbf{r}})/r]\exp[i\phi(\hat{\mathbf{r}})]$ where both A and ϕ are real-valued functions of direction and independent of r. In this case the surfaces of constant $\phi(\hat{\mathbf{r}})$ will not in general be spheres. It ultimately does not matter how one describes the geometry of the surfaces of constant phase; the crucial feature of the form of the far-field wave is that the scattering amplitude $\mathbf{E}_o^{sca}$ be independent of r.

field to produce a pattern of alternating radial inward and outward flow for $\langle \mathbf{S}^{cross} \rangle_t$. This happens because the relative phase between the waves changes with direction and distance from the particle and demonstrates that the waves interfere over *all directions*. This alternating flow structure is shown for points on $\mathcal{C}_h$ in Figs. (10.3) and (10.5) below, and one can see that the inward-outward flow becomes finer with angle as the distance R_{en} from the particle increases. The integration of these energy flows via Eq. (10.23) below, shown in Figs. (10.3) and (10.6), yields the extinction cross section C^{ext}. One can see that the energy flows in the side directions when integrated over a small angular range *nearly* yields zero due to the partial cancellation of neighboring regions of opposing radial flow. It is only when the angular range extends into the neighborhood of the forward direction that the partial cancellation becomes *more* incomplete, contributing to a nonzero net inward flow through $\mathcal{S}_{en}$ back toward the particle. Moreover, it is only after all directions, including the forward direction, are included in the integral that the full extinction cross section is obtained.

It is only in the forward direction that the relative phase between the incident and scattered waves is constant in r, and it is this behavior of the relative phase and the planar surfaces of constant phase of the incident wave that give the optical theorem its unidirectional character. One can see the independence of the relative phase between these waves in the forward direction by comparing the exponential terms in Eqs. (3.1), (3.2), (10.1) and (10.2). The independence of $\mathbf{E}_o^{sca}$ on r is what makes the relative phase constant in the forward direction and is a direct consequence of the far-field form of the scattered wave. This graphical-based reasoning can be made more rigorous. In the process, the role of absorption in extinction is revealed and a connection between the extinction cross section and a particle's physical characteristics is obtained.

In the following, the observation point is restricted to the $\mathcal{C}_h$ contour in the horizontal scattering plane as in Sec. 6.1.1. For points on $\mathcal{C}_h$

$$\hat{\mathbf{r}} \cdot \mathbf{E}_o^{inc} = \hat{\mathbf{r}} \cdot \mathbf{E}_o^{sca} = 0 \qquad \mathbf{r} \in \mathcal{C}_h, \tag{10.18}$$

and using Eqs. (3.1), (3.2), (10.1), (10.2) and (10.7), the radial component of the interference

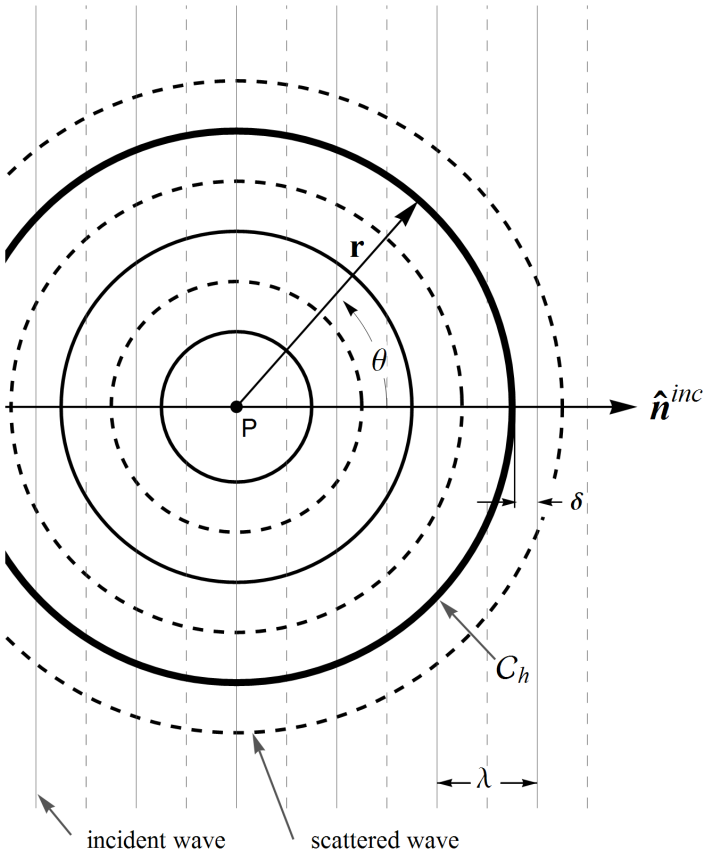

Figure 10.2: Interference of the incident and scattered waves. The surfaces of constant phase for the two waves are drawn separated by intervals of π in phase in the plane containing the origin and perpendicular to $\mathbf{E}_o^{inc}$. The light and dark lines correspond to the surfaces for the incident and scattered waves, respectively. The relative difference in phase between the waves in the forward direction is indicated by δ. Dashed and solid lines represent maxima and minima of the waves' amplitudes. Solid-solid (dash-solid) line intersections qualitatively indicate an outward (inward) radial component of the $\langle \mathbf{S}^{cross} \rangle_t$ energy flow, recall Eq. (10.19).

term in the total Poynting vector can be expressed as

$$\langle \mathbf{S}^{cross}(R_{en}\hat{\mathbf{r}})\rangle_t \cdot \hat{\mathbf{r}} = \frac{c_o}{R_{en}} K(\hat{\mathbf{r}}) \mathrm{Re}\Big\{ f_o(\hat{\mathbf{r}}) \exp\left[ikR_{en}(1-\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}^{inc})\right]\Big\} \qquad \mathbf{r}\in\mathcal{C}_h, \tag{10.19}$$

where $c_o = (\epsilon_o/\mu_o)^{1/2}$, and

$$K(\hat{\mathbf{r}}) = 1 + \hat{\mathbf{r}}\cdot\hat{\mathbf{n}}^{inc}\,, \tag{10.20}$$

and

$$f_o(\hat{\mathbf{r}}) = \mathbf{E}_o^{inc*}\cdot\mathbf{E}_o^{sca}(\hat{\mathbf{r}})\,. \tag{10.21}$$

Note that $\langle \mathbf{S}^{cross}\rangle_t$ has components along the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ directions also, but it is only the $\hat{\mathbf{r}}$ component that carries energy away from (or towards) the particle and hence is of interest here. The term K in Eq. (10.19) is the so-called obliquity factor which vanishes in the backward direction $-\hat{\mathbf{n}}^{inc}$ and originates from the counter-propagation of the two waves in that direction.

The significance of the two terms inside of the Re{...} filter in Eq. (10.19) requires some care to describe. The first term, f_o, is a complex-valued function of angle that accounts for the variation in magnitude and phase of the scattered electric field. By comparing Eqs. (10.3) and (10.21), one can see that f_o is independent of r and receives its angular functionality from the phase factor $\exp(-ik\,\mathbf{r}'\cdot\hat{\mathbf{r}})$ in the VIE, cf. Eq. (10.3). This shows that f_o describes the influence of the particle's physical properties, like size and shape, on the interference energy flow. By comparing the exponential factors appearing in Eq. (3.1) and (10.1), one finds that the factor $\exp[ikR_{en}(1-\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}^{inc})]$ in Eq. (10.19) accounts for the angular dependence of the *relative* difference in phase between the incident and scattered waves. This term has no dependence on the physical properties of the particle and hence accounts for the features of the interference energy flow that are a consequence of the form of the waves in the far-field zone.

The angular structure of $\langle \mathbf{S}^{cross}\rangle_t$ shown in Fig. (10.3) can now be explained. As $\mathbf{r}$ advances along the $\mathcal{C}_h$ contour the term $\exp[ikR_{en}(1-\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}^{inc})]$ oscillates, causing the radial component of $\langle \mathbf{S}^{cross}\rangle_t$ to alternate from being directed away from to toward the particle.

The function f_o is an angular envelope bounding this energy flow. Moreover, if one considers the intensity of the scattered wave, which is given by Eq. (3.28) as $I^{sca}(\hat{\mathbf{r}}) = r^2 \langle \mathbf{S}^{sca}(\mathbf{r}) \rangle_t \cdot \hat{\mathbf{r}}$, then Eqs. (10.1), (10.2) and (10.18) show that

$$I^{sca}(\hat{\mathbf{r}}) = \frac{c_o}{2} \mathbf{E}_o^{sca}(\hat{\mathbf{r}}) \cdot [\mathbf{E}_o^{sca}(\hat{\mathbf{r}})]^*, \tag{10.22}$$

which involves the same angular functionality as f_o, cf., Eq. (10.21). Figure 10.3(b) shows I^{sca} and K superimposed on the interference flow. Both curves are normalized to the forward direction and one can see the influence that I^{sca} and K have on the angular shape of the overall envelope bounding the flow.

10.5 Example of the Energy Flow for Single Particles

Simulations of spherical particles are examined below to illustrate how $\langle \mathbf{S}^{cross} \rangle_t$ describes extinction and integrates over $\mathcal{S}_{en}$ to yield C^{ext}. Both the scattered fields and the extinction cross section for the particle are found from Mie theory.

First consider a small spherical particle with size parameter $kR = 4.0$ and refractive index $m = 1.10 + 0i$, where R is the sphere radius. The propagation and polarization directions of the incident wave are taken to be $\hat{\mathbf{n}}^{inc} = \hat{\mathbf{z}}$ and $\mathbf{E}_o^{inc} = \hat{\mathbf{x}}$, respectively. The polar plots in Fig. 10.3(a)-10.3(c) show the radial component of the $\langle \mathbf{S}^{cross} \rangle_t$ energy flow along the $\mathcal{C}_h$ contour for successively larger values of R_{en}. To see how this energy flow integrates over $\mathcal{S}_{en}$ to yield C^{ext} through Eq. (10.12), consider the integral

$$\partial C^{ext}(\theta_s) = \frac{1}{I^{inc}} \int_{\partial \mathcal{S}_b} \langle \mathbf{S}^{cross}(\mathbf{r}) \rangle_t \cdot \hat{\mathbf{r}} \, \mathrm{d}S \,. \tag{10.23}$$

The symbol ∂ is not intended to represent a derivative. The open surface $\partial \mathcal{S}_b$ in Eq. (10.23) is the part of the large (closed) spherical surface $\mathcal{S}_{en}$ that extends from $\theta = \pi$, the backward direction, to $\theta = \theta_s$, see Fig. 10.4(a). When $\theta_s = \pi$, $\partial \mathcal{S}_b$ disappears, and when $\theta_s = 0$, $\partial \mathcal{S}_b = \mathcal{S}_{en}$. Notice that there is a sign difference between Eqs. (10.12) and (10.23) so that

$$C^{ext} = -\partial C^{ext}(0) \,. \tag{10.24}$$

Shown in the right-hand column of Fig. (10.3) are plots of the integral $\partial C^{ext}/C^{ext}$ as θ_s varies from π down to zero. The value of C^{ext} is calculated directly from the Mie series to verify that the integral in Eq. (10.23) does indeed yield the correct cross section when $\theta_s = 0$.

In the polar plot in Fig. 10.3(a), $R_{en} = 10R$ and the $\langle \mathbf{S}^{cross} \rangle_t$ energy flow is seen to alternate from radially outward to inward as expected from Eq. (10.19). The zero line is labeled in this plot. Examination of the corresponding plot of $\partial C^{ext}/C^{ext}$ shows that it oscillates about zero until around $\theta_s = 0$, where $\partial C^{ext}(0)/C^{ext} = -1$, recall Eq. (10.24). This negative value at $\theta_s = 0$, where the partial surface $\partial \mathcal{S}_b$ becomes complete coinciding with $\mathcal{S}_{en}$, indicates that $\langle \mathbf{S}^{cross} \rangle_t$ *carries a net inward flow of energy through* $\mathcal{S}_{en}$. This *net* inward flow *is* the extinction and demonstrates why the negative sign appears in front of the integral in Eq. (10.12), ensuring that $C^{ext} \geqslant 0$.

Figures 10.3(b) and 10.3(c) show the energy flow for increasing $\mathcal{C}_h$ contour radii $R_{en} = 20R$ and $R_{en} = 100R$, respectively. These polar plots demonstrate that the flow becomes finer with increasing R_{en}, but remains unchanged in the forward and backward directions as expected from Fig. (10.2). One can also clearly see the diminishment of the flow near the backward direction.

Now consider the two pairs of peaks labeled **d**, **e** and **f**, **g**. Comparison of these peaks shows that each successive peak is slightly greater in magnitude than the peak preceding it in θ_s. Moreover, the difference in magnitude between successive peaks increases as θ_s approaches zero. Consequently, the partial cancellation of angular regions of opposing energy flow becomes less complete nearer to the forward direction. In the limit that $kR_{en} \to \infty$, the peaks would be spaced infinitesimally close together in angle and would be infinite in number but would still occur in opposing-flow peak-pairs. In this case the contribution to the extinction cross section would still originate from the (infinitesimally) incomplete cancellation of neighboring angular regions of opposing flow. This limiting behavior of the energy flow is the physical significance behind the method of stationary phase as it is used

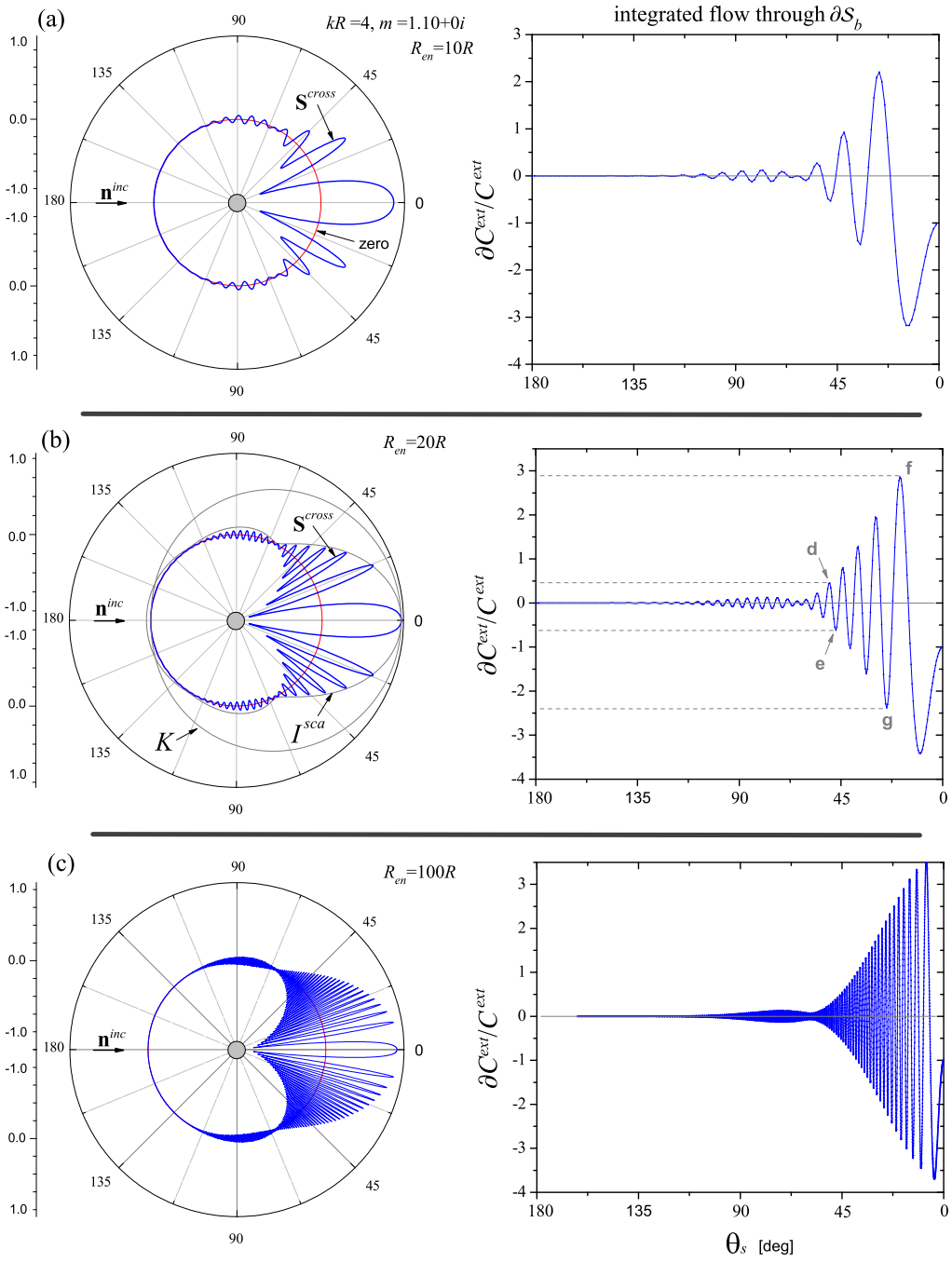

Figure 10.3: **Left column:** Polar plots of $\langle \mathbf{S}^{cross} \rangle_t \cdot \hat{\mathbf{r}}$ on the $\mathcal{C}_h$ contour as a function of angle in the plane perpendicular to the polarization of the incident field. The curves are normalized to the forward direction. The particle is a sphere with $kR = 4.0$ and $m = 1.10 + 0i$. The plots labeled (a), (b) and (c) show the energy flow for increasing $\mathcal{C}_h$ contour radii $R_{en} = 10R$, $R_{en} = 20R$ and $R_{en} = 100R$, respectively. Zero is indicated in polar plot (a). **Right column:** Plots of the integral given by Eq. (10.23) normalized by C^{ext} for the contour radii indicated in the corresponding polar plots.

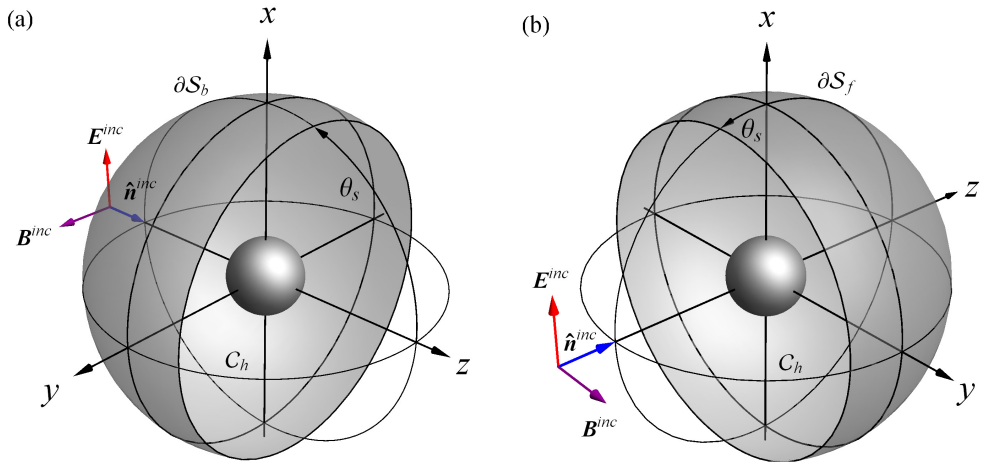

Figure 10.4: The partial surfaces $\partial\mathcal{S}_b$ and $\partial\mathcal{S}_f$ surrounding the spherical particles used to generate the plots of $\partial C^{ext}/C^{ext}$ in Figs. (10.3) and (10.6) via Eq. (10.23). In sketch (a), the surface begins in the backward direction at $\theta = \pi$ and extends to $\theta = \theta_s$. In sketch (b), the surface begins in the forward direction at $\theta = 0$ and extends to $\theta = \theta_s$.

to derive the optical theorem, cf. [24, App. 12].

The plots of $\langle \mathbf{S}^{cross} \rangle_t$ in Fig. (10.3) share a key common feature. *The flow alternates radially more rapidly with angle as R_{en} increases, but the direction of the flow along $\hat{\mathbf{n}}^{inc}$ remains unchanged.* Inspection of these plots shows that when the integral in Eq. (10.12) adds up this alternating flow over $\mathcal{S}_{en}$, it will add together regions of opposing flow of *nearly* equal magnitude in the side scattering directions. These regions of opposing flow will partially cancel each other's contribution to the integral. The plots of $\partial C^{ext}/C^{ext}$ in Fig. (10.3) shows the cancellation between adjacent regions of opposing flow becoming less complete closer to the forward direction. *This is why extinction has the appearance of being a forward angle effect.*

To examine the affect of absorption on extinction, Fig. (10.5) shows polar-plots like Fig. (10.3), but for larger spheres with increasing values of Im(m). The plots show the same qualitative behavior as the nonabsorbent spheres, demonstrating that absorption does not remove the radially alternating character of the interference energy flow. Consequently, extinction is still achieved via the same interference mechanisms described above.

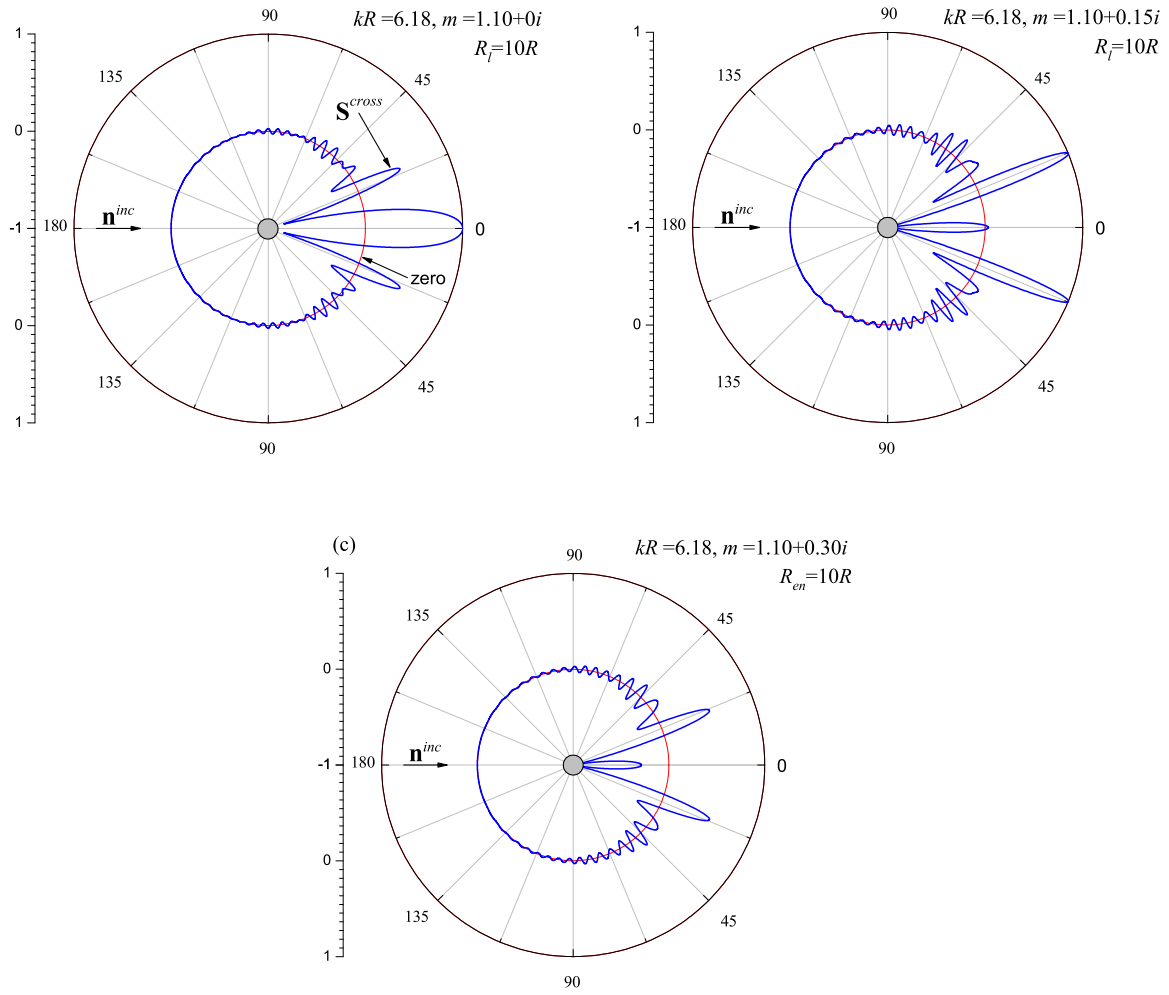

Figure 10.5: Polar plots of $\langle \mathbf{S}^{cross} \rangle_t \cdot \hat{\mathbf{r}}$ on the $\mathcal{C}_h$ contour like Fig. (10.3) except for spheres with varying degrees of absorption. The size parameter of the particles is $kR = 6.18$ and their refractive index varies from $m = 1.10 + 0i$ in (a), to $m = 1.10 + 0.15i$ in (b), to $m = 1.10 + 0.30i$ in (c). The contour radius is $R_{en} = 10R$.

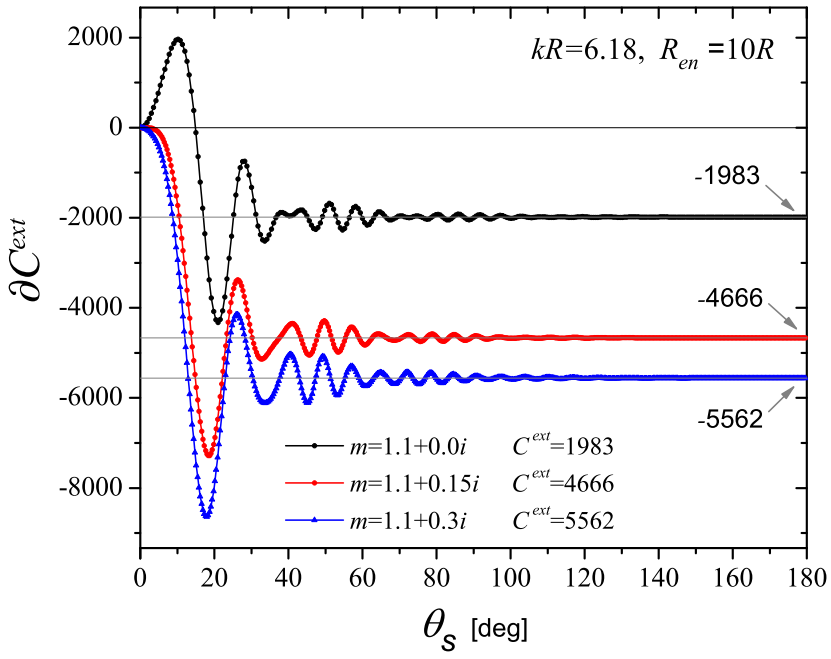

Figure 10.6: Plots of the integral $\partial C^{ext}/C^{ext}$ of Eq. (10.23) for the interference energy flow due to the spheres in Fig. (10.5). The partial surface $\partial\mathcal{S}_f$ used to generate the curves extends from $\theta = 0$ to $\theta = \theta_s$ and is shown in Fig. 10.4(b). The size of the surface is $R_{en} = 20R$. The values for C^{ext} in the legend are calculated directly from the Mie series for comparison to the values shown on the curves at $\theta_s = \pi$.

Notice that between Figs. (10.3) and (10.5), the increased size and refractive index of the particles causes the direction of energy flow along $\hat{\mathbf{n}}^{inc}$ to change. This unexpected behavior is explained by evaluating Eq. (10.19) in the $\hat{\mathbf{n}}^{inc}$ direction with the substitution of Eqs. (5.6) and (10.28) below, giving

$$\langle \mathbf{S}^{cross}(r\hat{\mathbf{n}}^{inc})\rangle_t \cdot \hat{\mathbf{n}}^{inc} = 2\,A\,\frac{c_o}{r}\,\cos\sigma. \tag{10.25}$$

Because both C^{ext} and A are necessarily non-negative, Eq. (10.29) below shows that σ above is bounded in $[0,\pi)$. Equation (10.25) then reveals that the energy flow *should* alternate in direction along $\hat{\mathbf{n}}^{inc}$ as σ varies with changes in the physical properties of the particle through Eq. (10.28).

Figure (10.6) shows the integral of Eq. (10.23) for the absorbent spheres in Fig. (10.5), except for a different perspective, the integral is taken over the partial surface $\partial\mathcal{S}_f$ shown in Fig. 10.4(b). This surface begins in the forward direction at $\theta = 0$ and extends to $\theta = \theta_s$. Indicated in the legend of Fig. (10.6) are the extinction cross sections for each of the three spheres as calculated directly from the Mie series. Also indicated on each curve is the value of ∂C^{ext} when the partial surface $\partial\mathcal{S}_f$ closes at $\theta_s = \pi$ and coincides with $\mathcal{S}_{en}$. The use of $\partial\mathcal{S}_f$ instead of $\partial\mathcal{S}_b$ demonstrates the expected result that it does not matter from which direction, forward or backward, the alternating energy flow is integrated. The curves clearly show that the dominant contribution to the cross section occurs near the forward direction as stated in [47, p. 31]. Comparing Figs. (10.5) and (10.6) demonstrates that *the direction of the energy flow along* $\hat{\mathbf{n}}^{inc}$ *can be either toward or away from the particle and the correct extinction cross section is still obtained upon integration of the flow over all directions.*

10.6 Connection to the Particle's Physical Properties

Using the VIE for the scattering amplitude, it is possible to demonstrate how the physical properties of the particle, i.e., its size, shape and refractive index, affect the extinction cross section. From Eq. (10.3), the amplitude profile f_o appearing as the argument of the $\mathrm{Im}\{\ldots\}$

filter in the optical theorem [Eq. (10.17)] is

$$\mathbf{E}_o^{inc*} \cdot \mathbf{E}_o^{sca}(\hat{\mathbf{n}}^{inc}) = \frac{k^2}{4\pi}\left(m^2 - 1\right) \int_V \mathbf{E}_o^{inc*} \cdot \mathbf{E}^{int}(\mathbf{r}') \, \exp(-ik\,\mathbf{r}' \cdot \hat{\mathbf{n}}^{inc}) \, \mathrm{d}\mathbf{r}', \tag{10.26}$$

where use has been made of the transverse character of the far-field scattered wave and the relation $\mathbf{E}_o^{inc*} \cdot (\hat{\mathbf{n}}^{inc} \otimes \hat{\mathbf{n}}^{inc}) \cdot \mathbf{E}^{int} = 0$. Using Eq. (10.26) one can define an amplitude and phase factor as

$$A = \left| \mathbf{E}_o^{inc*} \cdot \mathbf{E}_o^{sca}(\hat{\mathbf{n}}^{inc}) \right| , \tag{10.27}$$

and

$$\sigma = \tan^{-1} \left\{ \frac{\mathrm{Im}\left[\mathbf{E}_o^{inc*} \cdot \mathbf{E}_o^{sca}(\hat{\mathbf{n}}^{inc})\right]}{\mathrm{Re}\left[\mathbf{E}_o^{inc*} \cdot \mathbf{E}_o^{sca}(\hat{\mathbf{n}}^{inc})\right]} \right\} , \tag{10.28}$$

respectively, which when combined with Eq. (10.17) gives

$$C^{ext} = \frac{4\pi}{k|\mathbf{E}_o^{inc}|^2} A \, \sin\sigma \, . \tag{10.29}$$

This result is essentially equivalent to the form appearing in [47, 98].

To see how the optical theorem relates the particle's physical properties to C^{ext} through Eq. (10.29), first suppose that the particle is absorbent. This would mean that $\mathrm{Im}(m) \geqslant 0$. Referring to Eq. (10.28) shows that this can cause σ to increase, and thus increase C^{ext}, as compared to an identical particle without absorption. This provides a connection between the particle's absorption and its influence on the extinction cross section; but there is more to this. The appearance of absorption in a particle causes its internal field to change as compared to its nonabsorbent counterpart. With absorption, the field magnitude decays with distance into the particle, and near its inner surface, the field becomes directed more tangential to the surface, see [15, p. 352-356]. This then provides an indirect connection and could counteract the enhancement in C^{ext} due to $\mathrm{Im}\{m\}$ as mentioned above. Similarly, the size of the particle, its shape and its orientation if it is nonspherical, all directly affect the structure of the internal field and hence the value of the integral appearing in Eqs. (10.27) and (10.28) leading to the cross section through Eq. (10.29).

10.7 Implications for Measurement

Perhaps the most intuitive way to measure the extinction cross section is as the difference in the response of a detector facing into the forward direction with and without a particle present. This measurement is illustrated in Fig. (10.1) and is discussed in detail in [47, Ch. 4]. The essential idea behind the measurement is that the total power that reaches the detector's face is reduced by the presence of the particle, and that the reduction is proportional to C^{ext}. The following examines this measurement within the context of the energy flow plots presented in Figs. (10.3) and (10.6).

The response of a detector is determines by the component of the *total* energy flow $\langle \mathbf{S} \rangle_t$ directed into and integrated over its sensitive face. When the particle is present in the beam, $\langle \mathbf{S} \rangle_t$ is given by Eq. (10.4) and contains three contributions: $\langle \mathbf{S}^{inc} \rangle_t$, $\langle \mathbf{S}^{sca} \rangle_t$, and $\langle \mathbf{S}^{cross} \rangle_t$. Each of these contributions have different dependencies on the distance r between the particle and the detector. From Eqs. (10.1), (10.2) and (10.5)-(10.7) one can see that $\langle \mathbf{S}^{inc} \rangle_t$ is independent of r, $\langle \mathbf{S}^{sca} \rangle_t$ decays as r^{-2}, and $\langle \mathbf{S}^{cross} \rangle_t$ decays as r^{-1}. Consequently, in the far-field zone the total energy flow $\langle \mathbf{S} \rangle_t$ will be dominated by the contributions from $\langle \mathbf{S}^{inc} \rangle_t$ and $\langle \mathbf{S}^{cross} \rangle_t$.

Now consider Fig. (10.7) which shows the same situation as in Fig. 10.1(b) except with the three contributions to $\langle \mathbf{S} \rangle_t$ explicitly indicated along with their relative magnitude. Here the detector is the partial spherical surface $\partial \mathcal{S}_f$ shown in Fig. 10.4(b) with an angular size given by θ_s. The detector is located in the particle's far-field zone and its size is taken to be much greater than the particle's geometric projection into the forward direction but small enough as not to exceed the width w of the incident beam.

The component of the total energy flow due to $\langle \mathbf{S}^{inc} \rangle_t$ is approximately constant over the detector's face and is directed into it. This means that the detector receives, in part, an amount of power equal to the product of the flux of the incident wave I^{inc} and the area of its face. Added to this power is the contribution due to $\langle \mathbf{S}^{cross} \rangle_t$ which is given by ∂C^{ext} in Eq. (10.23). The magnitude of the contribution from $\langle \mathbf{S}^{inc} \rangle_t$ is always greater than that

of $\langle \mathbf{S}^{cross} \rangle_t$ in the far-field zone due to the latter's inverse dependence on r. Consequently, the detector always encounters a *net* flow of energy directed *into* its face and hence always registers a positive power. However, the contribution to this net power due to $\langle \mathbf{S}^{cross} \rangle_t$ can either subtract from or add to the contribution due to $\langle \mathbf{S}^{inc} \rangle_t$ depending on the detector's angular size. This is evident from Fig. (10.6) which shows that ∂C^{ext} can be either positive or negative depending on the value of θ_s. *It is only after a sufficiently large solid angle of the total energy flow is accommodated by the detector that the contribution from* $\langle \mathbf{S}^{cross} \rangle_t$ *reduces the power received by the detector by an amount corresponding to* C^{ext}. For the three particles considered in Fig. (10.6), the detector's angular size would have to be $\theta_s \geqslant 60^\circ$ in order to include enough of the forward interference energy flow to measure C^{ext} to within a few percent error.

Refer again to the plot in Fig. (10.6) and suppose that the angular size of the detector depicted in Fig. (10.7) is $\theta_s = 15^\circ$. Inspection of this plot shows that the contribution to the net power received by the detector due to $\langle \mathbf{S}^{cross} \rangle_t$ would be negative for the $m = 1.10 + 0.15i$ and $m = 1.10 + 0.30i$ particles whereas the it would be positive for the $m = 1.10 + 0i$ particle. This means that the power received by the detector when the $m = 1.10 + 0.15i$ or $m = 1.10 + 0.30i$ particle is present will be less than the power received by the detector in the particle's absence, as expected. However, if the $m = 1.10 + 0i$ particle is present, the power received by the detector is *greater* than when the particle is absent. Recalling that the curves in Fig. (10.6) are calculated from Fig. (10.5), one can see that the origin of this unexpected behavior is that the $\langle \mathbf{S}^{cross} \rangle_t$ energy flow near the forward direction is oriented away from the particle in Fig. 10.5(a) whereas it is oriented toward the particle in Figs. 10.5(b) and 10.5(c). Energy conservation is not violated in this example since the detector is not collecting all 4π steradians of the total energy flow, recall Eqs. (10.9)-(10.13).

The reader familiar with the use of Beer's law to measure the extinction coefficient of a dilute colloid or aerosol may be troubled that the measurements discussed above require a detector with an angular size exceeding 60°. Such measurements on dilute colloids and

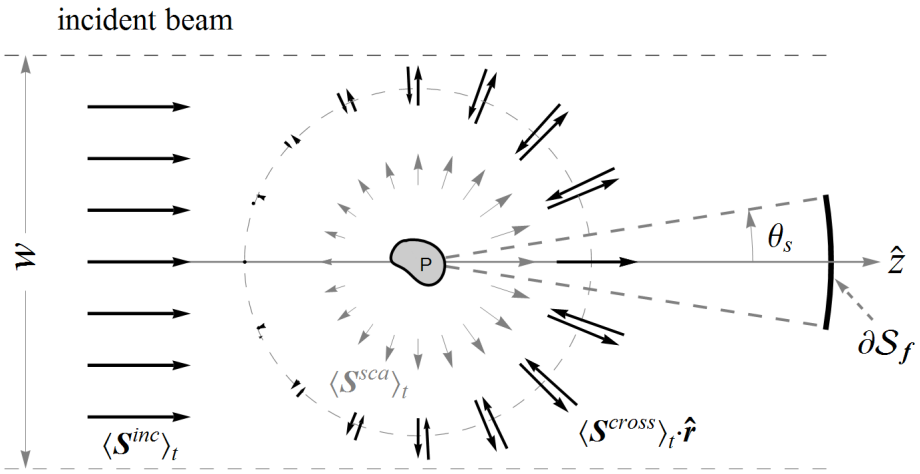

Figure 10.7: Same situation as in Fig. 10.1(b) except with the three contributions to the total energy flow qualitatively indicated. Here the detector is the partial surface $\partial\mathcal{S}_f$ shown in Fig. 10.4(b) and the solid angle subtended by the detector is given by θ_s.

aerosols are known to yield the correct extinction coefficient for far smaller detector solid angles. The remainder of this chapter will generalize the single-particle analysis presented above to multi-particle groups and demonstrate that the presence of multiple particles causes the dominant contribution to the group's extinction cross section to reside within a narrower angular region around the forward direction, and hence will resolve the apparent inconsistency with Beer's law.

10.8 Particle Group in the Noninteracting Limit

Here a group of identical particles will be examined in the non-interacting particle approximation. In reality, there is always some degree of interaction between the particles and the effect of this interaction will be addressed in Sec. 10.10. By considering non-interacting particles first, one is most easily able to compare to the results presented above involving single particles. The conditions required to minimize the interactions between the particles in a group are presented in [1, Ch. 7], [41, 101], [102, p. 2]. These conditions are reviewed here for completeness and to illustrate the limits of validity of the approximations used.

Consider a group of N identical particles residing in vacuum. The distribution of the particles is fixed in time, but otherwise arbitrary. Let R_{sc} be the radius of the smallest circumscribing sphere that encloses any one of the N particles. The origin of the laboratory coordinate system is located near the geometric center of a particle and this particle will be referred to as the primary particle $i = 1$, see Fig. 10.8(a). The location of the i^{th} particle is given by the vector $\mathbf{R}_i$ which connects the origin of the primary particle $\mathcal{O}_1$ to that of the i^{th} particle, $\mathcal{O}_i$. Let R_{gr} be the radius of the smallest circumscribing sphere that encloses the entire N-particle group. The observation point $\mathbf{r}$ is assumed to be located far enough from the group that it satisfies the far-field conditions of [43] with respect to *both* the group and each constituent particle.

Let R_{min} be the minimum inter-particle separation, see Fig. 10.8(b). The particles will be considered non-interacting if

$$k(R_{min} - R_{sc}) \gg 1, \tag{10.30}$$

$$R_{min} \gg R_{sc}, \tag{10.31}$$

$$kR_{min} \gg \frac{k^2 R_{sc}^2}{2}. \tag{10.32}$$

The conditions that establish that the observation point is in the far-field zone of the entire group are given by Eqs. (4.13)-(4.15) as

$$k(r - R_{gr}) \gg 1, \tag{10.33}$$

$$r \gg R_{gr}, \tag{10.34}$$

$$kr \gg \frac{k^2 R_{gr}^2}{2}. \tag{10.35}$$

The inequalities of Eqs. (10.30)-(10.35) are the criteria for the far-field single scattering approximation (SSA), where the entire group effectively acts as a single point-like particle [1, Ch. 7], [101].

A consequence of the identical particle SSA is that the far-field scattered wave due to the group can be formulated in terms of the wave scattered by the primary particle only.

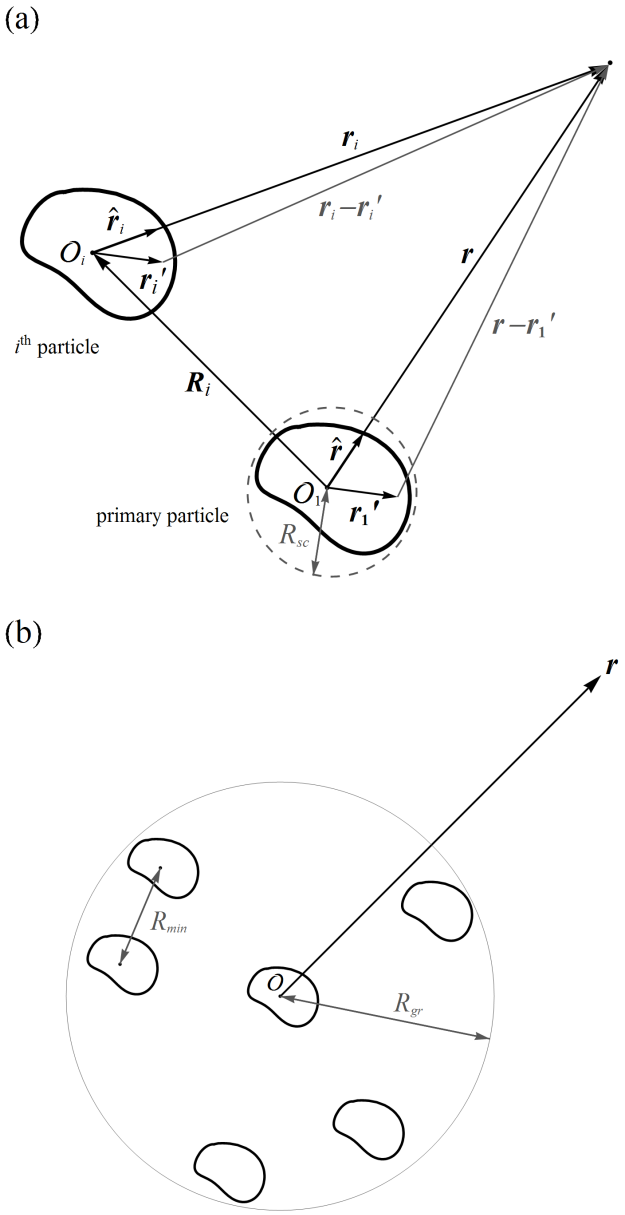

Figure 10.8: Sketch (a) shows the primary particle and the i^{th} particle of a multi-particle group. Sketch (b) illustrates the definition of the particle group circumscribing-sphere-radius R_{gr} and the minimum inter-particle separation R_{min}.

To describe how this is done, denote the primary particle as $i = 1$ and consider the volume integral equation (VIE) giving the scattered electric field for *any* external location,

$$\mathbf{E}_1^{sca}(\mathbf{r}) = k^2\left(m^2-1\right)\int_{V_1}\overleftrightarrow{G}(\mathbf{r},\mathbf{r}_1')\cdot\mathbf{E}_1^{int}(\mathbf{r}_1')\,\mathrm{d}\mathbf{r}_1'\,. \tag{10.36}$$

Here V_1 and $\mathbf{E}_1^{int}$ denote the volume and internal electric field of the primary particle, respectively. Recall that Eq. (10.36) is valid at any point $\mathbf{r}$, not just in the far-field zone.

The task now is to determine how the electric field inside of the i^{th} particle ($i \neq 1$) is related to $\mathbf{E}_1^{int}$. This can be done using the dyadic transition operator, which relates the incident field inside of a particle to its scattered wave as [1, p. 117]

$$\mathbf{E}_1^{sca}(\mathbf{r}) = \int_{V_1}\overleftrightarrow{G}(\mathbf{r},\mathbf{r}_1')\cdot\int_{V_1}\overleftrightarrow{T}(\mathbf{r}_1',\mathbf{r}_1'')\cdot\mathbf{E}^{inc}(\mathbf{r}_1'')\,\mathrm{d}\mathbf{r}_1''\,\mathrm{d}\mathbf{r}_1' \qquad \mathbf{r}\in\mathbb{R}^3, \tag{10.37}$$

where

$$\overleftrightarrow{T}(\mathbf{r}_1,\mathbf{r}_1') = k^2\left(m^2-1\right)\delta(\mathbf{r}_1-\mathbf{r}_1')\overleftrightarrow{I} + k^2\left(m^2-1\right)\int_{V_1}\overleftrightarrow{G}(\mathbf{r}_1,\mathbf{r}_1'')\cdot\overleftrightarrow{T}(\mathbf{r}_1'',\mathbf{r}_1')\,\mathrm{d}\mathbf{r}_1'' \qquad \mathbf{r}_1,\mathbf{r}_1'\in V_1. \tag{10.38}$$

Comparing Eqs. (10.36) and (10.37) shows that the primary particle's internal field obeys

$$k^2\left(m^2-1\right)\mathbf{E}_1^{int}(\mathbf{r}_1) = \int_{V_1}\overleftrightarrow{T}(\mathbf{r}_1,\mathbf{r}_1')\cdot\mathbf{E}^{inc}(\mathbf{r}_1')\,\mathrm{d}\mathbf{r}_1' \qquad \mathbf{r}_1\in V_1. \tag{10.39}$$

From Eq. (10.38), one can see that $\overleftrightarrow{T}$ depends on the particle's material properties through m, the vacuum wavelength through k, and the particle's shape and size through the integral over V_1. Notice however, that $\overleftrightarrow{T}$ must be independent of the choice of origin since the functions δ and $\overleftrightarrow{G}$ in Eq. (10.38) do not depend on any specific origin. This gives the translation behavior of the operator as

$$\overleftrightarrow{T}(\mathbf{r}_1,\mathbf{r}_1') = \overleftrightarrow{T}(\mathbf{r}_i,\mathbf{r}_i'), \tag{10.40}$$

where $\mathbf{r}_i = \mathbf{r}-\mathbf{R}_i$ and $\mathbf{r}_i' = \mathbf{r}_1'+\mathbf{R}_i$, recall Fig. 10.8(a). Using Eq. (3.1) and Eqs. (10.39), (10.40) and Fig. 10.8(b), the internal electric field at a location $\mathbf{r}_i'$ inside of the i^{th} particle is related to the internal field at the corresponding point $\mathbf{r}_1'$ inside of the primary particle as

$$\mathbf{E}^{int}(\mathbf{r}_i') = \exp(ik\mathbf{R}_i\cdot\hat{\mathbf{n}}^{inc})\,\mathbf{E}^{int}(\mathbf{r}_1') \quad \mathbf{r}_i'\in V_i,\ \mathbf{r}_1'\in V_1. \tag{10.41}$$

Now consider the simplification of the dyadic Green's function appearing in Eq. (10.36) for the i^{th} particle in the SSA. Equation (10.34) justifies

$$\frac{1}{|\mathbf{r}_i - \mathbf{r}_i'|} \approx \frac{1}{r}, \tag{10.42}$$

and

$$\hat{\mathbf{r}}_i \approx \hat{\mathbf{r}}, \tag{10.43}$$

whereas Eqs. (10.33)-(10.35) together permit

$$\exp(ik|\mathbf{r}_i - \mathbf{r}_i'|) \approx \exp[ik(r - \mathbf{r}_i' \cdot \hat{\mathbf{r}})]. \tag{10.44}$$

With the approximations of Eqs. (10.42)-(10.44), one can simplify the dyadic Green function to

$$\overleftrightarrow{G}(\mathbf{r}, \mathbf{r}_i') \approx \frac{1}{4\pi}\frac{e^{ikr}}{r}\left(\overleftrightarrow{I} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}\right)\exp(-ik\mathbf{r}_i' \cdot \hat{\mathbf{r}}), \tag{10.45}$$

which from Eq. (10.36), gives the scattered electric field due to the i^{th} particle as

$$\mathbf{E}_i^{sca}(\mathbf{r}) = \frac{k^2}{4\pi}\frac{e^{ikr}}{r}\left(m^2 - 1\right)\left(\overleftrightarrow{I} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}\right) \cdot \int_{V_i} \mathbf{E}_i^{int}(\mathbf{r}_i') \exp(-ik\mathbf{r}_i' \cdot \hat{\mathbf{r}})\, \mathrm{d}\mathbf{r}_i'\,. \tag{10.46}$$

With the use of Eq. (10.41) and $\mathbf{r}_i' = \mathbf{r}_1' + \mathbf{R}_i$, the integral over the i^{th} particle's volume V_i in Eq. (10.46) can now be replaced by the integral over the primary particle's volume V_1, giving

$$\mathbf{E}_i^{sca}(\mathbf{r}) = \frac{k^2}{4\pi}\frac{e^{ikr}}{r}\left(m^2 - 1\right)\exp[ik\mathbf{R}_i \cdot (\hat{\mathbf{n}}^{inc} - \hat{\mathbf{r}})]\left(\overleftrightarrow{I} - \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}\right) \cdot \int_{V_1} \mathbf{E}_1^{int}(\mathbf{r}_1') \exp(-ik\mathbf{r}_1' \cdot \hat{\mathbf{r}})\, \mathrm{d}\mathbf{r}_1'\,. \tag{10.47}$$

Equation (10.47) is equivalent to the far-field scattered electric field given by Eqs. (4.17) and (10.3) augmented by the phase factor $\exp[ik\mathbf{R}_i \cdot (\hat{\mathbf{n}}^{inc} - \hat{\mathbf{r}})]$. Recall the scattering wave vector $\mathbf{q}$ from Eq. (7.2)

$$\mathbf{q}(\hat{\mathbf{r}}) = k(\hat{\mathbf{n}}^{inc} - \hat{\mathbf{r}}),$$

then, Eq. (10.47) becomes

$$\mathbf{E}_i^{sca}(\mathbf{r}) = \exp\left[i\mathbf{q}(\hat{\mathbf{r}}) \cdot \mathbf{R}_i\right]\, \mathbf{E}_1^{sca}(\mathbf{r}), \tag{10.48}$$

which gives the scattered electric field of the i^{th} particle in terms of that of the primary particle. The net wave scattered by the entire non-interacting particle group is then given by

$$\mathbf{E}_{gr}^{sca}(\mathbf{r}) = \frac{\exp(ikr)}{r}\mathbf{E}_{o}^{sca}(\hat{\mathbf{r}})\sum_{i=1}^{N}\exp\left[i\mathbf{q}(\hat{\mathbf{r}})\cdot\mathbf{R}_i\right], \tag{10.49}$$

$$\mathbf{B}_{gr}^{sca}(\mathbf{r}) = \frac{k}{\omega}\frac{\exp(ikr)}{r}\,\hat{\mathbf{r}}\times\mathbf{E}_{o}^{sca}(\hat{\mathbf{r}})\sum_{i=1}^{N}\exp\left[i\mathbf{q}(\hat{\mathbf{r}})\cdot\mathbf{R}_i\right], \tag{10.50}$$

where $\mathbf{E}_{o}^{sca}$ is the single-particle scattering amplitude of Eq. (10.3). The interference term in the Poynting vector that describes the extinction due to the particle group follows from Eqs. (3.1), (3.2), (10.49), and (10.50) as

$$\langle\mathbf{S}_{gr}^{cross}(\mathbf{r})\rangle_t = \frac{1}{2\mu_o}\,\mathrm{Re}\Big\{\mathbf{E}^{inc}(\mathbf{r})\times\left[\mathbf{B}_{gr}^{sca}(\mathbf{r})\right]^{*} + \mathbf{E}_{gr}^{sca}(\mathbf{r})\times\left[\mathbf{B}^{inc}(\mathbf{r})\right]^{*}\Big\}. \tag{10.51}$$

Following the same analysis as in Sec. 10.5, the structure of $\langle\mathbf{S}_{gr}^{cross}\rangle_t$ is examined by surrounding the group by the large imaginary spherical surface $\mathcal{S}_{en}$ of radius R_{en} centered on the origin. Remember that the intersection of $\mathcal{S}_{en}$ with the plane containing the origin and perpendicular to the polarization of the incident electric field defines the horizontal scattering plane and the $\mathcal{C}_h$ contour, recall Fig. (6.2). Integration of the radial component of the interference energy flow of Eq. (10.51) over $\mathcal{S}_{en}$ and normalization by the intensity of the incident wave I^{inc}, gives the group's extinction cross section as

$$C_{gr}^{ext} = -\frac{1}{I^{inc}}\oint\limits_{\mathcal{S}_{en}}\langle\mathbf{S}_{gr}^{cross}(\mathbf{r})\rangle_t\cdot\hat{\mathbf{r}}\,\mathrm{d}S\,. \tag{10.52}$$

This is a direct generalization of Eqs. (10.7) and (10.12), which give the single particle cross section C^{ext}. Application of the optical theorem, Eq. (10.17), to Eq. (10.49) above shows that

$$C_{gr}^{ext} = N\,C^{ext}. \tag{10.53}$$

Equation (10.53) is the classic result that the total extinction cross section C_{gr}^{ext} for a group of non-interacting particles is the sum of each particle's cross sections independently, see [47, Sec. 4.22] and [1, Sec. 7.2]. Notice that the calculations above demonstrate

that the additivity of the cross sections *does not* rely on any assumption that the spatial distribution of the particles be random, as is assumed in [47, Sec. 4.22]. The form of Eqs. (10.49) and (10.50) shows that the particle group's scattered wave is a transverse spherical wave in the far-field zone. Consequently, the same interference mechanism that enables the redistribution of energy in the extinction caused by a single particle applies to a non-interacting multi-particle group. The only difference between single and multi-particle extinction in this case is the appearance of the phase-factor sum in Eqs. (10.49) and (10.50).

Using Eqs. (3.1), (3.2), (10.49), (10.50), and (10.52) above, one finds that the radial component of the interference energy flow through the $\mathcal{C}_h$ contour is

$$\langle \mathbf{S}_{gr}^{cross}(R_{en}\hat{\mathbf{r}})\rangle_t \cdot \hat{\mathbf{r}} = \frac{c_o}{R_{en}} K(\hat{\mathbf{r}}) \operatorname{Re}\Big\{ f_o(\hat{\mathbf{r}})\, f_{gr}(\hat{\mathbf{r}}) \exp\left[ikR_{en}(1-\hat{\mathbf{r}}\cdot\hat{\mathbf{n}}^{inc})\right]\Big\} \qquad \mathbf{r}\in\mathcal{C}_h, \quad (10.54)$$

where

$$f_{gr}(\hat{\mathbf{r}}) = \sum_{i=1}^{N} \exp\left[i\mathbf{q}(\hat{\mathbf{r}})\cdot\mathbf{R}_i\right]. \quad (10.55)$$

When compared to the corresponding single-particle energy flow of Eq. (10.19), Eq. (10.54) reveals that the spatial distribution of the particles in the group introduces an additional envelope bounding the alternating radial energy flow, but is otherwise identical.

10.9 Examples for a Non-Interacting Group

To illustrate the similarities and differences between single and multi-particle extinction in the SSA, simulations of the interference energy flow for a multi-particle group consisting of identical spheres are examined below. The size and refractive index of each sphere is $kR = 6.18$ and $m = 1.25 + 0.30i$ where R is the (single) sphere radius. The particles are distributed within the group volume in both an ordered and disordered manner. The ordered distribution has the particles centered on a cubic lattice of spacing $R_o = 4R$ that fills a spherical volume of radius R_{gr} and is shown in Fig. 10.9(a) where $N = 136$. The disordered distribution is shown in Fig. 10.9(b) and is generated by displacing the particles in the ordered distribution by a random distance between $-R$ and R along each coordinate

axis. The disordered distribution is not intended to be rigorously random in character; rather, its purpose is to demonstrate what effect the structure of the distribution has on the group's extinction. Notice that the inter-particle separation of $4R$ for these distributions does not satisfy Eqs. (10.30)-(10.32), yet the particles are treated as non-interacting anyway. The reason that this is done is due to the excessive computational demands[3] resulting from an attempt to strictly satisfy both Eqs. (10.30)-(10.32) and Eqs. (10.33)-(10.35) simultaneously. Care is taken, however, to satisfy Eqs. (10.33)-(10.35).

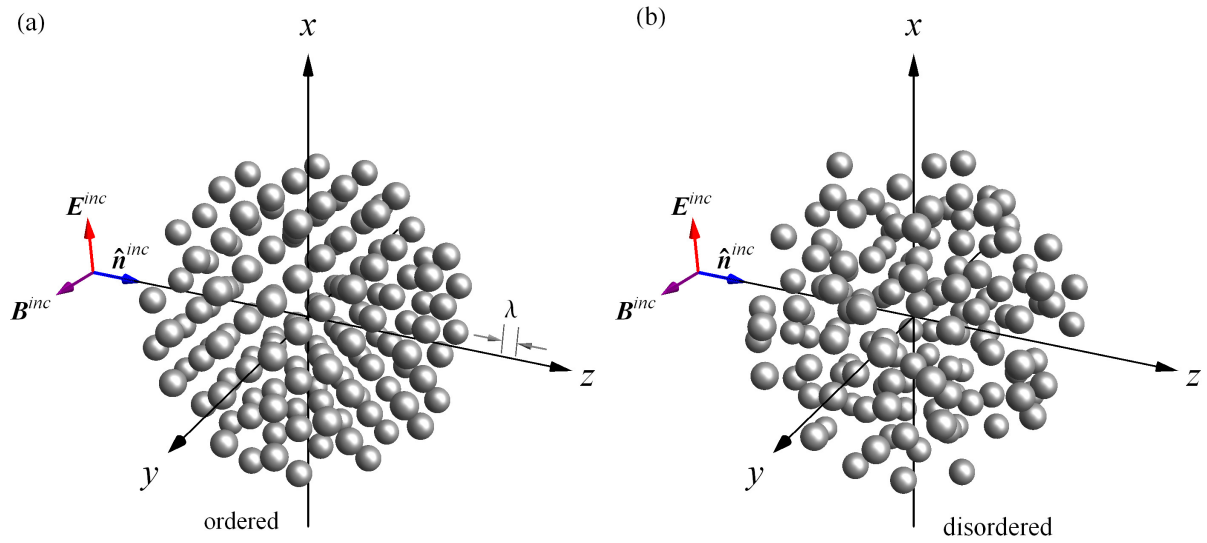

Figure 10.9: Ordered (a), and disordered (b), spherical particle distributions used in the simulations of Fig. (10.10). The total number of particles in both groups is $N = 136$ and the vacuum wavelength λ is shown in (a) along the z-axis for scale.

The extinction due to the two spherical-particle groups is calculated using the Mie solution to the Maxwell equations for a single sphere. The net scattered fields for the entire multi-particle group are then obtained from the single particle fields through Eqs. (10.49) and (10.50). The interference energy flow due to the group is then calculated from Eq. (10.54).

[3]At least from with regard to the DDA. One *could* easily do this exactly for spherical particles using the superposition T-Matrix method.

Consider the analogy to Eq. (10.23) for the multi-particle group,

$$\partial C_{gr}^{ext}(\theta_s) = \frac{1}{I^{inc}} \int_{\partial \mathcal{S}_f} \langle \mathbf{S}_{gr}^{cross}(\mathbf{r}) \rangle_t \cdot \hat{\mathbf{r}} \, \mathrm{d}S \,, \tag{10.56}$$

where the surface $\partial \mathcal{S}_f$ is the part of the large spherical surface $\mathcal{S}_{en}$ extending from the forward direction $\theta = 0$ to θ_s as shown in sketch (b) of Fig. (10.4). Comparison of Eqs. (10.52) and (10.56) shows that $\partial C_{gr}^{ext}(\pi) = -NC^{ext}$ when $\theta_s = \pi$ and the partial surface $\partial \mathcal{S}_f$ closes to coincide with $\mathcal{S}_{en}$.

Plots of Eq. (10.56) are shown in Fig. (10.10) for the ordered and disordered spherical particle groups consisting of $N = 1$, 19, and 136 particles. The curves are normalized by the single-particle extinction cross section C^{ext} and the size of the partial surface $\partial \mathcal{S}_f$ used to render the curves is $R_{en} = 600 R_{gr}$. These plots demonstrate the same oscillating behavior as the single particle simulations in Fig. (10.3). Small detailed differences in the structure of the curves for the ordered and disordered distributions can be found, yet the curves achieve the value for C_{gr}^{ext} expected from Eq. (10.53) regardless of the distribution's structure. Moreover, the curves show that as the number of particles in the group increases, and hence the group size increases, the dominant contribution to the extinction cross section occurs over a decreasing angular region near the forward direction. The origin of this behavior is explained by the group amplitude profile f_{gr} in Eq. (10.55), which has the form of a Fourier series and hence narrows in $\mathbf{q}$ as the group grows in extent $\mathbf{R}_i$. More explicitly, f_{gr} of Eq. (10.55) can be transformed into a Fourier integral as

$$\sum_{i=1}^{N} \exp\left[i\mathbf{q}(\hat{\mathbf{r}}) \cdot \mathbf{R}_i\right] \rightarrow \frac{1}{V_{gr}} \int_{V_{gr}} \exp\left[i\mathbf{q}(\hat{\mathbf{r}}) \cdot \mathbf{r}'\right] \, \mathrm{d}\mathbf{r}'. \tag{10.57}$$

This transformation will be valid when $\mathbf{q}(\hat{\mathbf{r}}) \cdot \mathbf{d} \ll 1$, where $\mathbf{d}$ is the maximum nearest-neighbor separation between the particles in the group. From Eq. (7.2) one can see that this condition essentially restricts $\hat{\mathbf{r}}$ to point near the forward direction $\hat{\mathbf{n}}^{inc}$.

The integral in Eq. (10.57) can be approximately evaluated in terms of the radius of gyration of the particle group following Sec. 8.1. The result is

$$f_{gr}(\hat{\mathbf{r}}) \approx 1 - \frac{1}{3}\left[q(\hat{\mathbf{r}})R_g\right]^2 \quad \text{for} \quad q(\hat{\mathbf{r}})R_g < 1, \tag{10.58}$$

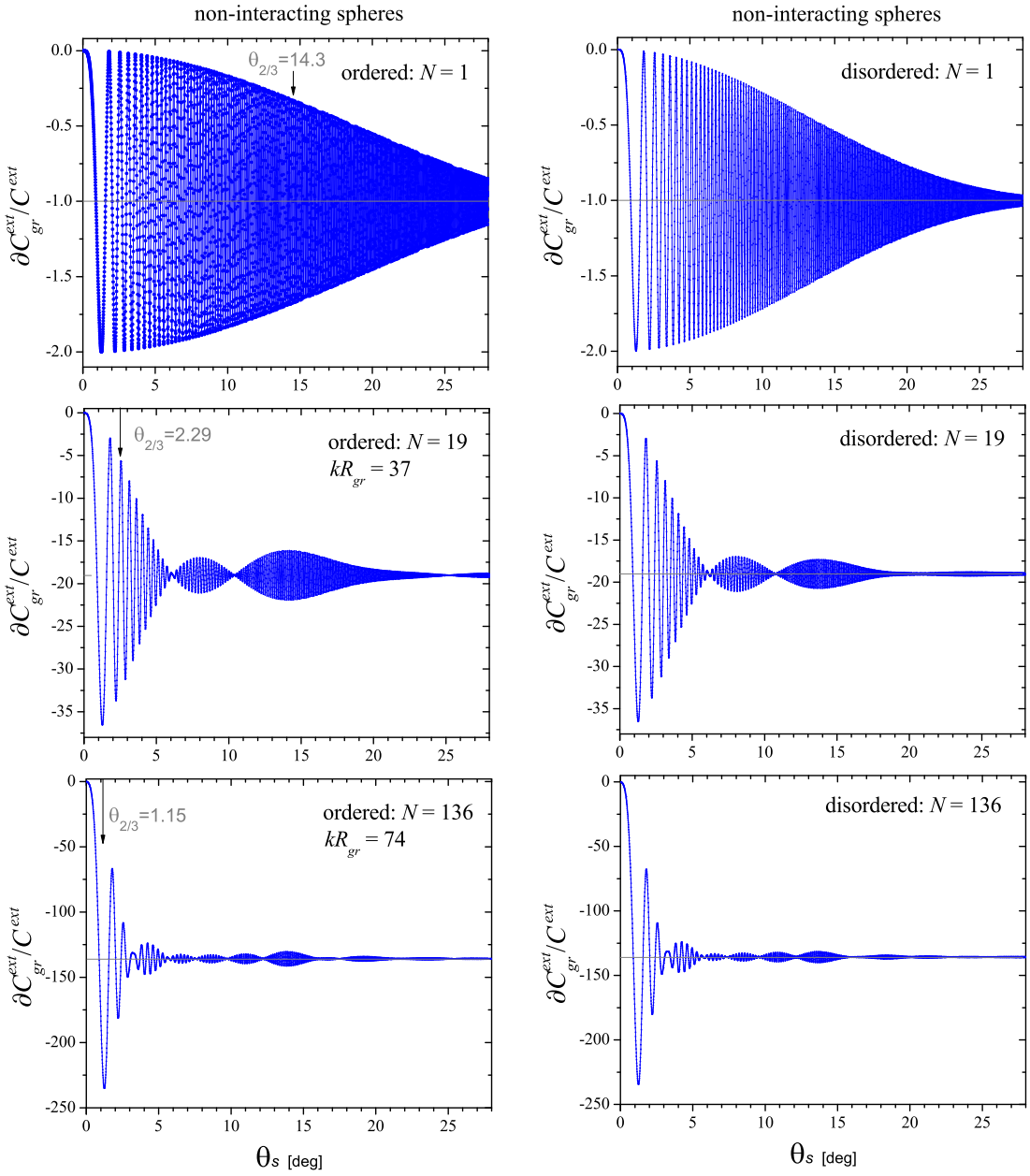

Figure 10.10: Plots of the integral ∂C_{gr}^{ext} given in Eq. (10.56) for the non-interacting ordered and disordered spherical particle groups. In each case the curves are normalized to the extinction cross section of a single particle C^{ext}. The size parameter and refractive index of a single particle is $kR = 6.18$ and $m = 1.25 + 0.30i$, respectively.

where R_g is the group's radius of gyration. If it is now assumed that the particle group is spherical and that the distribution of the particles appears approximately uniform on the scale of $1/q$ throughout the group volume, the group's radius of gyration becomes $R_g = \sqrt{3/5}R$, see [80]. Provided that the observation point resides on the $\mathcal{C}_h$ contour near the forward direction, one can use Eqs. (7.2) and (10.58) to find the approximate angle $\theta_{2/3}$ at which the magnitude of the interference energy flow of Eq. (10.51) decreases to two thirds of its value in the forward direction,

$$\theta_{2/3} \approx \sqrt{\frac{5}{3}}\frac{1}{kR_{gr}}\,. \tag{10.59}$$

The values of $\theta_{2/3}$ are indicated in Fig. (10.10) where the narrowing of the extinction around the forward direction with increasing group radius is evident.

10.10 Examples for an Interacting Group

Up to this point interactions between the particles of a group have been entirely neglected. Consequently, a strong similarity is found between the behavior of the energy flow due to a single particle and a non-interacting particle group. Here a small group of fully-interacting particles is examined using the DDA to see what effects the inter-particle interactions have on the extinction mechanism.

The fully-interacting particle group considered here is composed of $N = 9$ identical, homogeneous cubical particles distributed in a body-centered cubical arrangement, see Fig. (10.11). The sphere-volume-equivalent size parameter of the individual particles is $kR_{ve} = 3.38$ whereas that of the entire group is $kR_{ve}^{gr} = 7.02$. The particles' refractive index is $m = 1.33 + 0i$ and they all share the same orientation as shown in Fig. (10.11).

The DDA of Ch. 5 is used to simulate the group's scattered wave. This is done by using the DDA to numerically solve for the internal electric field inside of each particle complete with a full account of the inter-particle interactions. Once each particle's internal field is known, the group's scattered wave is calculated via Eq. (10.46) and the interference energy flow $\langle \mathbf{S}_{gr}^{cross}(\mathbf{r})\rangle_t$ follows from Eq. (10.51).

Plot (a) of Fig. (10.12) shows the integral ∂C_{gr}^{ext} given by Eq. (10.56) for the *single* cubical particle located at the origin in Fig. (10.11) in the absence of its eight neighbors, i.e. without interactions. The curve is normalized by the single-particle extinction cross section C^{ext} and the radius of the partial surface $\partial\mathcal{S}_f$ is $R_{en} = 10R_{gr}$. This plot is qualitatively identical to Fig. (10.6) and the $N = 1$ plot in Fig. (10.10). As before, one can see that the interference energy flow must be integrated on $\partial\mathcal{S}_f$ beyond 60° from the forward direction to obtain C^{ext} to better than 10% error.

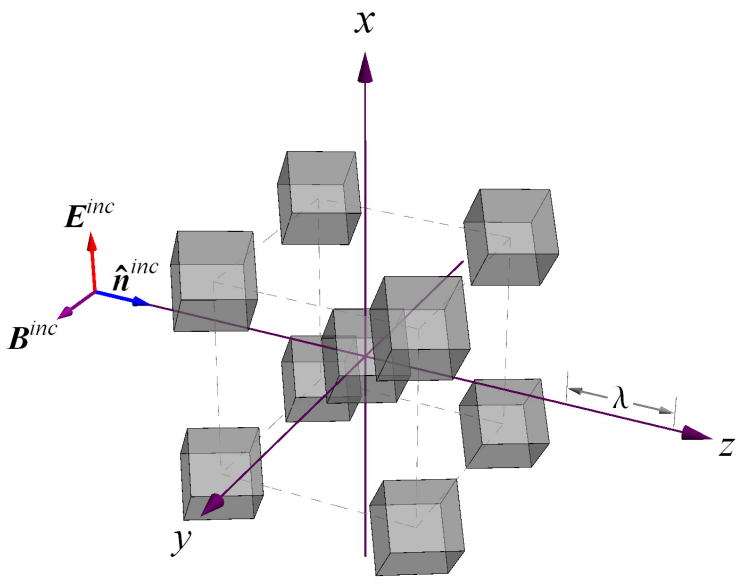

Figure 10.11: Body-centered cubical arrangement of identical cubical particles corresponding to the integrated energy flow plots shown in Figs. (10.12) and (10.13). The vacuum wavelength λ is shown for scale.

Figure 10.12 plot (b) shows the integrated energy flow ∂C_{gr}^{ext} for the entire interacting particle group. Here the curve is normalized by nine times the single particle extinction cross section. The qualitative effect of the presence of multiple particles is exactly the same as before; the dominant contribution to the group's extinction is compressed into a narrower angular region around the forward direction as compared to that of a single constituent particle. Inspection of plot (b) in Fig. 10.12 reveals that the curve converges to a value slightly greater than -1. This shows that the particle-group extinction cross section C_{gr}^{ext} is

greater than what would be expected in SSA from Eq. (10.53). Note that this increase is not necessarily a general feature of the effect of inter-particle interactions on the extinction cross section, see [1, Sec. 7.4] for more detail.

To further investigate the effect of inter-particle interactions, Fig. (10.13) shows ∂C_{gr}^{ext} normalized by $9C^{ext}$ for the group when the particles are treated as both non-interacting and interacting. Comparison of the two curves shows that the larger value of C_{gr}^{ext} for the interacting group comes from slight increases and decreases in the magnitude of the integrated energy flow. This indicates that here the primary influence of the inter-particle interactions occur mostly in the immediate neighborhood of the forward direction.

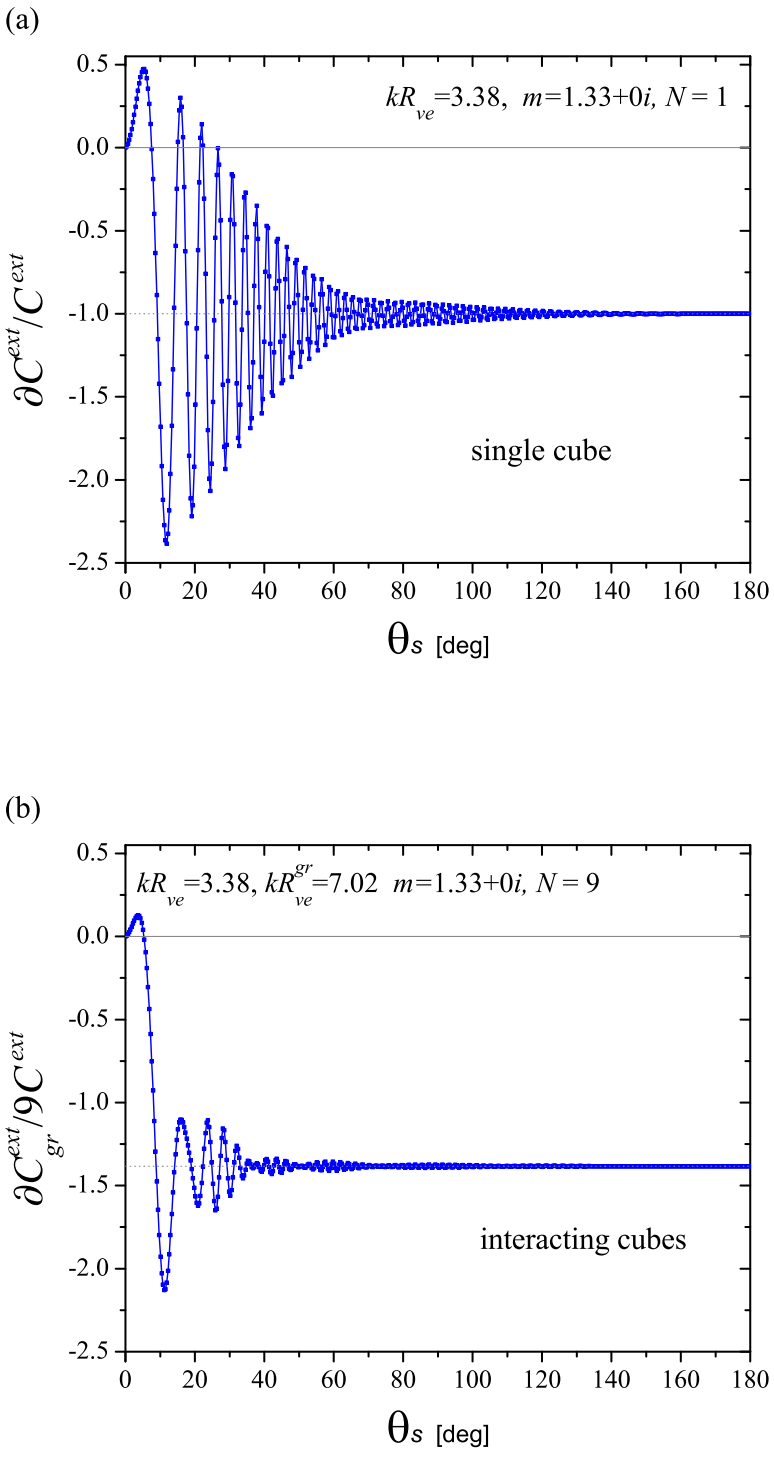

Figure 10.12: Plots of ∂C^{ext}_{gr} given in Eq. (10.56) for a single cube (a), and the interacting cubical particle group (b). In plot (a) the curve is normalized by the single-particle extinction cross section C^{ext} whereas the curve in plot (b) is normalized by $9C^{ext}$.

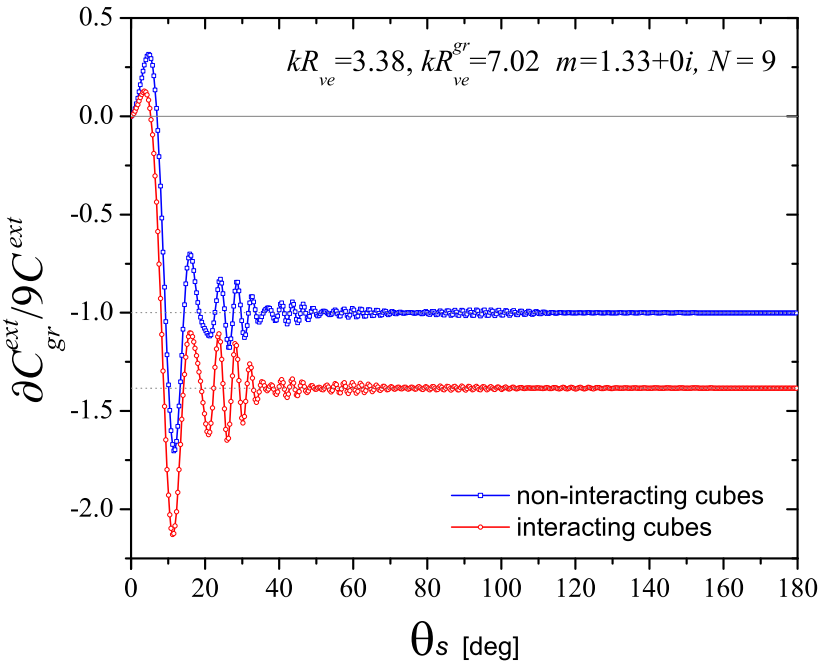

Figure 10.13: Plot of the integral ∂C_{gr}^{ext} given in Eq. (10.56) for the cubical particle group when the particles are treated as non-interacting and interacting. Both curves are normalized by $9C^{ext}$ where C^{ext} is the extinction cross section of a single cubical particle.

10.11 Discussion

An important conclusion from this chapter is that extinction does not necessarily result in the reduction of the power received by a detector facing into the forward direction when a particle, or particle group, is present. Whether or not the detector in the far-field zone receives a reduced amount of power depends on the detector's angular size. The magnitude of the total energy flow along the exact forward direction can either be reduced or enhanced as compared that of the incident wave.

Recall the RDG approximation used in Chs. 7 and 8, where a particle's internal field is replaced by the incident field. If one makes this substitution in Eq. (10.3) and feeds the result into Eq. (10.17), one finds that $C^{ext} = 0$ unless the refractive index has a nonzero imaginary part, i.e., the particle is absorbent. But, this cannot be correct since it says that such particles should have an extinction cross section that depends only on their absorptive character, hence energy conservation is violated. This issue is discussed frequently in the literature, where it is usually attributed to the inaccuracy of the forward scattering amplitude in the RDG approximation [90, p. 72]. One of the fruits of examining extinction from the microphysical perspective, is that one can now understand that this "inaccuracy" is caused by the lack of refraction in the RDG approximation. From Eqs. (10.26) and (10.28), one can see that a slight deviation in the phase of the internal field with respect to the incident field would result in an imaginary part for the forward scattering amplitude, even if $\mathrm{Im}\{m\} = 0$. Such a phase shift is caused by refraction, recall Sec. 5.6. This shows that *a particle's wavelet sources must have a nonzero phase shift relative to the incident wave in order to achieve the redistribution of energy associated with extinction.*

Recent work by [35] shows that the unexpectedly large detector solid angle that can be required in order to accurately measure C^{ext} appears to be exaggerated by the spherical shape of $\partial\mathcal{S}_f$ used above to integrate the energy flow. If the partial surface $\partial\mathcal{S}_f$ is replaced by a square surface of similar size, the integration of the interference energy flow results in curves like Fig. (10.3) and (10.10) that converge to the expected C^{ext} faster with increasing

solid angle. The reason that the spherical geometry for the detector appears to be a special case is because the surfaces of constant phase of the far-field scattered wave intercept the detector in circular contours that are exactly centered on and degenerate in shape with the detector. Recall Fig. (10.2) and sketch (b) in Fig. (10.4). When the detector radius increases, the phase of the scattered wave on the detector's edge advances uniformly over the entire detector-edge circumference. To compare this to the square detector case, one can think of the square as being separated into a circular surface and four "corner pieces." The phase of the scattered wave varies over the circular portion as the overall detector size increases exactly as it does for the circular detector. However, the corner pieces experience a different functionality of scattered-wave phase advance with increasing detector size. This adds extra oscillations to the energy flow that are not in "harmony" with the circular part, evidently resulting in a damping of the cross section integral, Eq. (10.23).

Chapter 11

Extinction Paradox

> *"The explanation of this apparent contradiction is that no matter how large the obstacle may be and no matter how far away from it the field is considered, there is always a narrow region - the neighborhood of the edge of the geometrical shadow - where the geometrical optics approximation does not hold."*
>
> -M. Born [24, p. 722]

One of the remarkable predictions of Mie theory is that the extinction cross section of a spherical particle asymptotically approaches a value of twice the area of its geometric shadow G as the size of the sphere becomes large relative to the wavelength. This extinction paradox is an unexpected result in the sense that geometrical optics should apply in the large-size regime, in which case the sphere[1] should extinct only as much power as is incident across G.

An intuitive explanation of the paradox, developed by van de Hulst, attributes one factor of G in the cross section to reflection, refraction, and absorption by a particle's illuminated profile, and the other factor to diffraction of the incident wave from the region on the particle's surface where its geometrical shadow begins [47, Sec. 8.2], [24, p. 722], [93, p. 68]. An alternative explanation, developed by [48] and later improved by [104], explains the paradox as being a consequence of the formation of a shadow immediately behind the particle. The highly intuitive character of these prevailing explanations has resulted in the

[1] Recent work has shown that the paradox also occurs for large particles of other shape [103].

extinction paradox being widely regarded as a well-understood phenomenon.

This chapter describes subtle problems with these popular explanations of the paradox and thus demonstrates a need for a more complete understanding. Specifically, the idea in the opening quote above will be shown to be plainly wrong. A novel view of extinction will be developed based upon the microphysical model. This approach facilitates a formal connection between the extinction cross section and the Ewald-Oseen extinction theorem. *The end result is that the paradox is ultimately seen to be a consequence of the requirement that the incident wave be canceled throughout a particle's interior.*

11.1 A Paradox

To describe the extinction paradox in detail, consider a uniform spherical particle residing in vacuum and illuminated by the plane wave of Eqs. (3.1) and (3.2). A spherical particle is chosen here for simplicity and because Mie theory can be used to calculate the fields exactly. This is only an example; *the primary conclusions of this work are not restricted to spherical particles.* The particle is centered at the coordinate origin, and enclosing the particle is the same imaginary spherical surface $\mathcal{S}_{en}$ introduced in Ch. 3. This surface is also centered on the origin and has an outward normal denoted $\hat{\mathbf{n}}_{en}$ and radius R_{en}, see Fig. (11.1).

As mentioned in Sec. 1.7, the presence of the particle in the incident wave establishes a new wave that can be described as the superposition of the incident and a scattered waves,

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{inc}(\mathbf{r}) + \mathbf{E}^{sca}(\mathbf{r}),$$

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}^{inc}(\mathbf{r}) + \mathbf{B}^{sca}(\mathbf{r}).$$

From the total fields, one finds C^{ext} from Eq. (3.5)

$$C^{ext} = -\frac{1}{2\mu_o I^{inc}} \operatorname{Re} \oint_{\mathcal{S}_{en}} \left\{ \mathbf{E}^{inc}(\mathbf{r}) \times \left[\mathbf{B}^{sca}(\mathbf{r})\right]^* + \mathbf{E}^{sca}(\mathbf{r}) \times \left[\mathbf{B}^{inc}(\mathbf{r})\right]^* \right\} \cdot \hat{\mathbf{r}} \, \mathrm{d}S,$$

where $I^{inc} = (1/2)\sqrt{\epsilon_o/\mu_o}|\mathbf{E}_o^{inc}|^2$ is the intensity of the incident wave and the asterisk denotes complex conjugation. Recall Eq. (3.5) and Sec. 10.2. Note that $\hat{\mathbf{r}} = \hat{\mathbf{n}}_{en}$ on $\mathcal{S}_{en}$ here. From

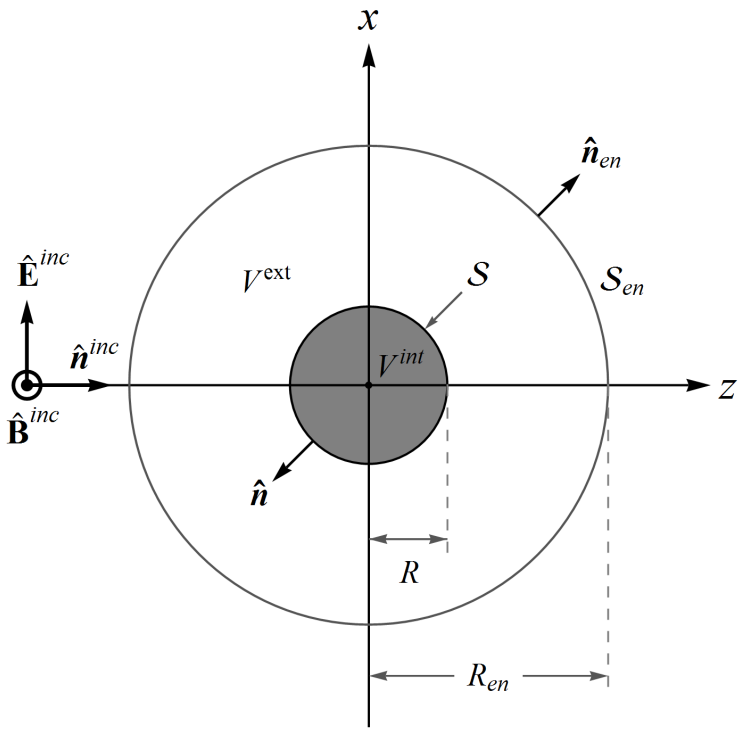

Figure 11.1: Scattering arrangement.

C^{ext}, the particle's extinction efficiency factor Q^{ext} follows as

$$Q^{ext} = \frac{C^{ext}}{G}, \tag{11.1}$$

where G is the geometrical projection of particle into the forward direction $\hat{\mathbf{n}}^{inc}$.

To explain the physical meaning of Q^{ext}, let the volume inside $\mathcal{S}_{en}$, including the particle, define the "system." As discussed in Sec. 10.2, conservation of energy shows that the power quantity $I^{inc}C^{ext}$ is equal to the net power lost by the system due to the outward-traveling component of the scattered wave and any absorption within the particle. *This net power loss constitutes an analytical definition of the extinction and does not necessarily associate extinction with a reduction of the energy flow in the forward direction.* Recall Sec. 10.7. Then, the meaning of Q^{ext} is the amount of power removed from the *system* due to scattering and absorption relative to the amount of power contained in the incident wave geometrically intercepted by the particle.

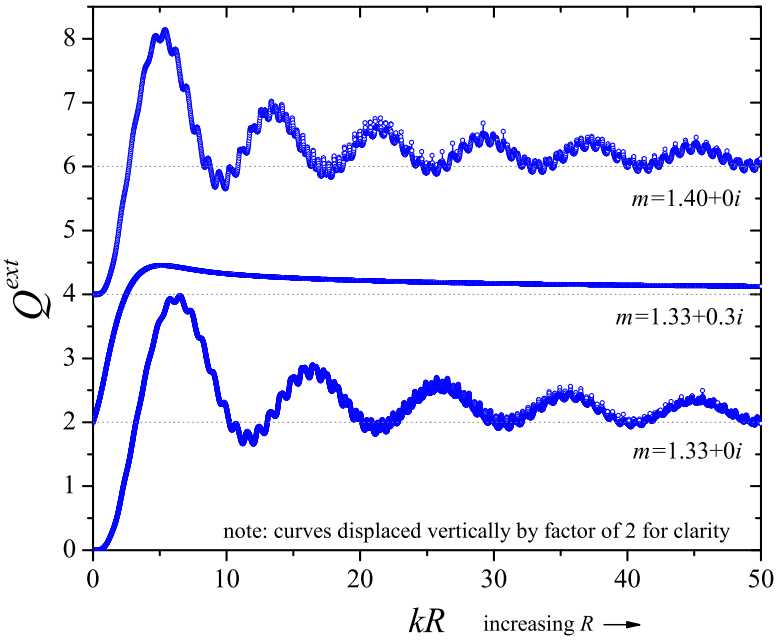

Figure 11.2: Extinction efficiency factor Q^{ext} for spherical particles as a function of increasing size kR, where k is constant for all curves. The refractive index of each curve is shown. The top two curves are displaced vertically by an additive factor of 2 for clarity.

The extinction efficiency factor for spheres of several refractive indices is shown in Fig. (11.2) as a function of kR, where k is fixed. One can see that each curve increases from zero, undergoes a succession of large and small-scale oscillations, and asymptotically approaches a value of $Q^{ext} = 2$ as $kR \to \infty$. These three regions of the curves are referred to as the Rayleigh, resonance, and asymptotic regions, respectively. Explanations for the behavior of Q^{ext} in the Rayleigh and resonance regions are given in multiple references including [37, Sec. 4] and [105].

The asymptotic behavior of Q^{ext} as $kR \to \infty$ is commonly known as the extinction paradox. This is because in the $kR \to \infty$ limit, one would expect that geometrical optics is valid. Then the particle should redirect through reflection, refraction, and absorption only the portion of light that it geometrically intercepts. Since these effects[2] all constitute losses to the energy contained in $\mathcal{S}_{en}$, i.e., the system, one would expect that $Q^{ext} = 1$. Hence

[2]Remember that refraction is a characteristic of the internal field, the latter of which ultimately leads to the form of the scattered wave.

$Q^{ext} = 2$ is paradoxical.

11.2 Popular Explanations of the Paradox

The most commonly encountered explanation of the paradox is due to van de Hulst [47]. This explanation, which will be called the "diffraction explanation" here, relies on the idea that geometrical optics fails in the shadow boundary where diffraction effects can become important. The shadow boundary is the region on a particle's surface where its geometrically illuminated side transitions to its shaded side, see Fig. (11.3). An alternative but less common explanation by Brillouin explains the paradox as a consequence of the incident wave being canceled within the geometrical shadow immediately behind the particle. This explanation will be called the "shadow explanation." These explanation are reviewed in the following[3] for completeness and to enable a detailed description of the subtle problems associated with them.

11.2.1 Diffraction Explanation

In the context of geometrical optics, the incident wave constitutes an infinitely wide beam of parallel rays along $\hat{\mathbf{n}}^{inc}$. For a large particle, these rays can be separated into two groups; those that intercept the particle's surface and those that do not, see Fig. (11.3). The rays that intercept the particle are reflected, refracted, and absorbed depending on m. Theses rays occupy a transverse area of the incident wavefront equal to G. The remaining rays, those not intercepted by the particle, form an incomplete wavefront with an area of G absent. The missing rays in this incomplete wave front cause the rest of the rays to diffract just as they would in passing by a large, opaque, transversely-oriented disk-like obstacle with the same shape as G, see Fig. (11.3).

From Babinet's principle, recall Sec. 9.3, one can understand that the combined amount

[3]One should note that the concepts of diffraction and rays are used below in the way in which they are intended in the original explanations. Often these views are not consistent with the understanding resulting from previous chapters.

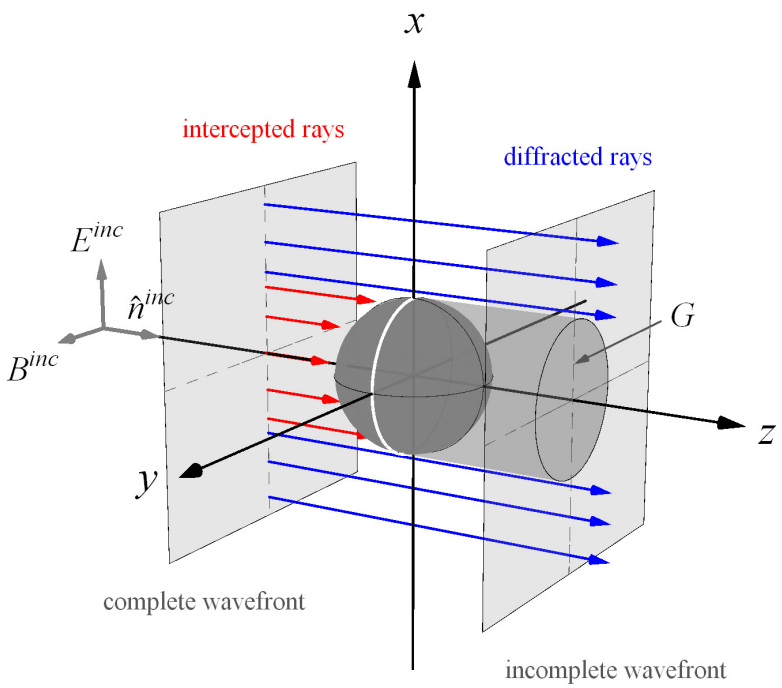

Figure 11.3: Arrangement of rays contained in the incident wave that are geometrically intercepted by the particle (red) and those that are not (blue). The particle's geometrical projection G in the froward direction is shown along with the shadow boundary indicated by the white line on the particle's surface.

of power contained in these diffracted rays is equal to $I^{inc}G$. The power contained in the intercepted rays is also equal to $I^{inc}G$. If one now regards extinction as the removal of energy flow from the forward direction[4] $\hat{\mathbf{n}}^{inc}$, then one concludes that the intercepted and diffracted rays each contribute a factor of G to a large particle's extinction cross section, resulting in $Q^{ext} = 2$. Hence the gist of the diffraction explanation is that one factor of G in C^{ext} originates from reflection, refraction, and absorption whereas the other factor of G originates from diffraction at the particle's shadow boundary.

11.2.2 Shadow Explanation

The shadow explanation of the paradox does not require any explicit consideration of diffraction. Rather, it relies on the idea that the incident wave must be canceled immediately *behind* the particle inside its geometrical shadow if the particle is perfectly opaque. This cancellation is achieved by the particle's scattered wave through destructive interference with the incident wave. To see how this explains the paradox, the scattered wave is separated into two parts; a reflection wave and a shadow-forming wave (SFW). The SFW is the part of the scattered wave that achieves cancellation of the incident wave *just behind* the particle. The reflection wave is then the remaining part of the scattered wave, see Fig. (11.4).

To say more, consider the case when an opaque particle is nonabsorbent, i.e., $\mathrm{Re}\{m\} \to \infty$, a reflector. Since the SFW cancels the incident within the particle's geometrical shadow, it must carry a quantity of power $I^{inc}G$ along the forward direction. Moreover, because the particle is nonabsorbent, no energy can be removed from the system. Therefore, one can conclude that the reflection wave must also carry a power of $I^{inc}G$ (now in all directions) in order to compensate for the missing power resulting from the cancellation of the incident wave. Since there is no absorption, the total scattering and extinction cross sections must be equivalent. Therefore, the scattered wave carries a *net* power of $2I^{inc}G$ through $\mathcal{S}_{en}$, and hence, the extinction cross section becomes $2G$, yielding $Q^{ext} = 2$. One should note that the shadow explanation need not be restricted to opaque or reflecting particles only. The SFW

[4]This *is* the common operational definition, but recall its inconsistency with Ch. 10.

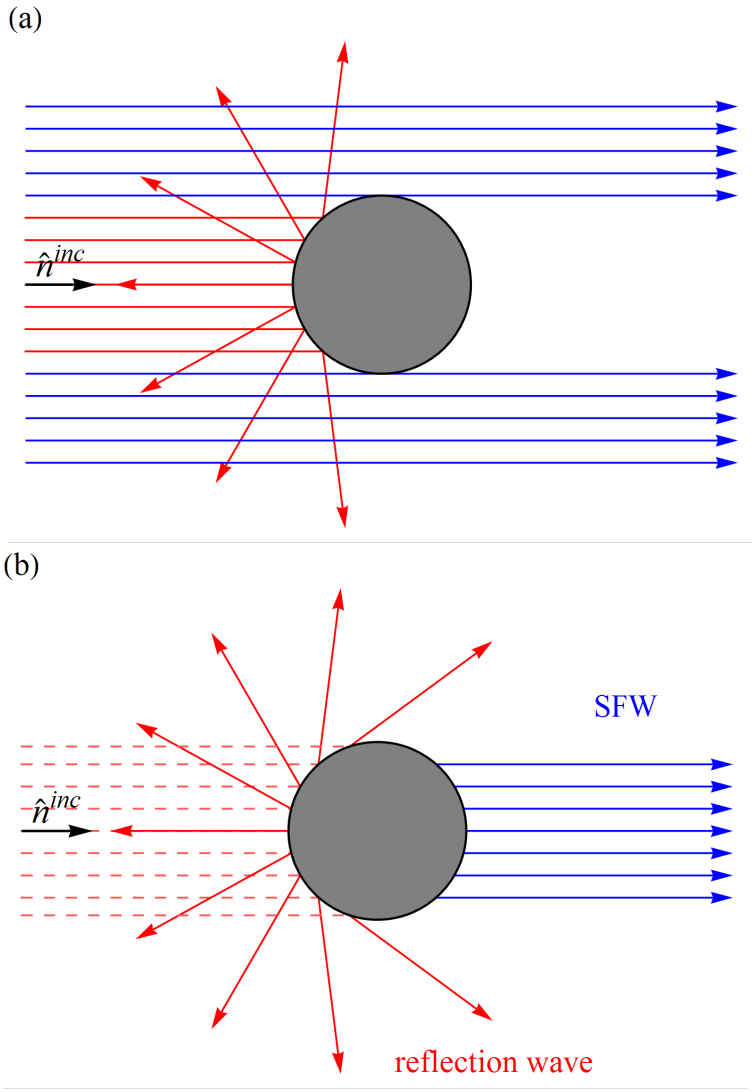

Figure 11.4: Shadow explanation of the extinction paradox. In (a) the vectors figuratively represent the total energy flow near a large perfectly conducting spherical particle. The red rays are intercepted by the particle's profile whereas the blue rays are not. Sketch (b) shows the energy flows associated with the SFW (blue) forming the particle's shadow in (a), and the energy flow of the reflected wave (red). The dashed lines indicate the incident wave's energy flow.

concept is generalized in [104] to apply to dielectric two-dimensional particles illuminated by a Gaussian beam.

11.3 Problems with the Popular Explanations

A fundamental assumption involved in the diffraction explanation is that the incident wave can be meaningfully separated into rays that are intercepted and diffracted by the particle. The implicit criteria validating this separation is that the particle's transverse dimensions be much greater than the vacuum wavelength. Consequently, the diffraction explanation really can apply only for the $kR \to \infty$ behavior of Q^{ext}. *However, one finds that the paradox also occurs when the particle size and wavelength are constant while the real part of the refractive index is increased.* With regard to spherical particles, the overall behavior of Q^{ext} for vastly different values of kR and m is qualitatively universal in terms of the phase shift parameter ρ given by Eq. (7.3). Recall that the physical meaning of ρ is the net difference in phase between a ray passing through the center of a spherical particle along a diameter and a ray in the incident wave traversing the same distance in vacuum.

Figure (11.5) shows Q^{ext} as a function of ρ when ρ is increased by increasing R *and* by increasing $\mathrm{Re}\{m\}$ as indicated. One can see that $Q^{ext} \to 2$ even for the rather small particle $kR = 10$. For a particle of such a size, and especially smaller, the separation of the incident wave into rays intercepted and rays diffracted by the particle *is not physically meaningful* [47, Sec. 3.3]. Only in the large size regime can the energy flow around, and within, a particle be described by collections of rays, which undergo rectilinear propagation, reflection, refraction, and absorption at the particle surface. For example, Fig. (11.6) shows the *total* energy flow for a sphere that is roughly λ wide, with $kR = 3$ and $m = 4.0 + 0i$, hence $\rho = 18$. This flow is calculated from Eq. (1.45). One can see that the flow cannot be described by collections of rays, unlike plot (a) in Fig. (11.14) below where this is more reasonable.

The other fundamental assumption of the diffraction explanation is that the particle's

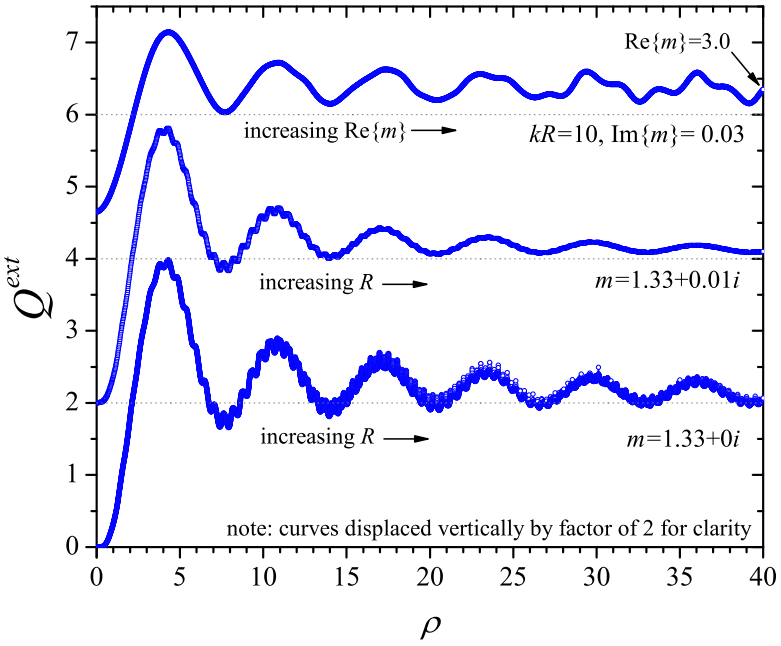

Figure 11.5: Extinction efficiency factor Q^{ext} for spheres as a function of ρ, Eq. (7.3). The phase shift is increased as indicated on the curves and the wavenumber is constant for all curves. Note that the top two curves are displaced vertically by a factor of two for clarity.

extinction cross section is calculated, or measured, in the far-field zone. It is only in the far-field that diffraction diverts enough energy flow from the forward direction as to account for a total factor of power $I^{inc}G$, given the operational definition. The problem is that this argument fails in the near-field. As one moves infinitesimally close to the particle's surface, i.e., by shrinking $\mathcal{S}_{en}$ to $\mathcal{S}$, the diversion of energy flow from the forward direction due to diffraction vanishes. Yet, it is shown in Ch. 3 that that the extinction cross section is completely independent of the distance from the particle. One must then find that $Q^{ext} \to 2$ even very close (compared to λ) to the particle surface, which is exactly where the concept of Fraunhoffer diffraction becomes invalid.

The separation of the incident wave's rays into those intercepted and those diffracted tacitly assumes that the particle has a well-defined surface in the sense of geometrical optics. This is because the ray separation requires that there is a clear distinction between the geometrically illuminated and shaded sides of the particle. It is possible to imagine situations

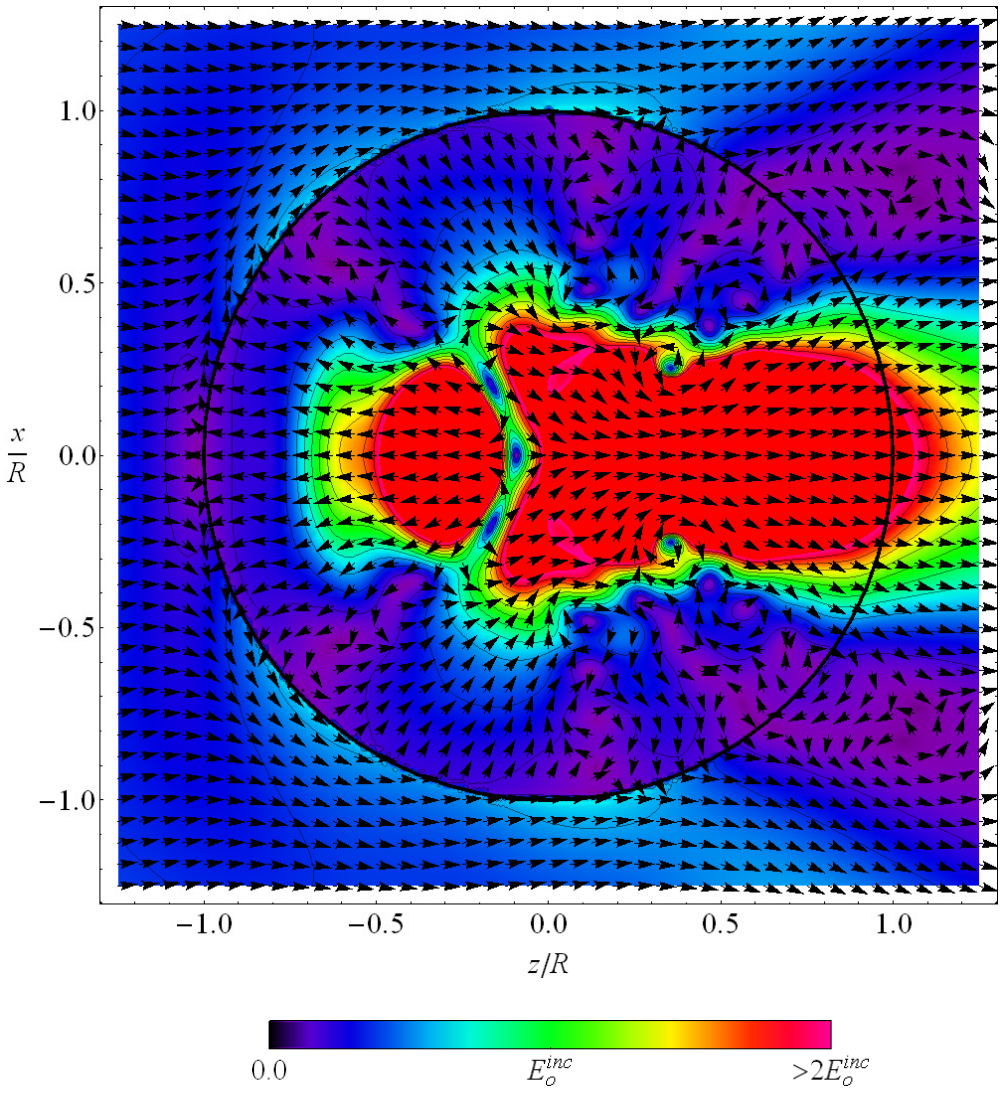

Figure 11.6: Time averaged energy flow, as given by the Poynting vector, inside a spherical particle and in its near-field. Here $kR = 3.0$ and $m = 4.0 + 0i$. The plot shows the energy flow in the x-z plane passing through the particle's center, recall Fig. (1.1). The color scale indicates the flow magnitude relative to the magnitude of the incident wave E_o^{inc}.

for which identification of such a shadow boundary is ambiguous. For example, consider a large particle with a radial refractive index profile that tails off to one exponentially with distance from the particle center. Or, consider a large fractal-like aggregate of small particles, like in Ch. 8. In both of these examples it is not clear where one would choose to identify a shadow boundary, or even a particle surface[5].

Now consider the shadow explanation of the paradox. Here there is no problem regarding an implied restriction to the far-field zone. Actually, the very concept of incident-wave cancellation in the shadow is realized in the extreme near-field just behind the particle. The shadow explanation is typically presented in the context of large, perfectly opaque particles[6]. However, as can be understood from Fig. (11.5), the extinction paradox occurs in the $\rho \to \infty$ limit, and hence concerns both opaque and dielectric particles equally. If the particle is dielectric the total wave just behind it in its geometric shadow is not necessarily zero and can be quite large due to focusing effects, e.g., see Fig. (11.6). In such a case, it becomes unclear why it is necessary to identify a SFW since there is no longer a requirement that the total wave vanish within the geometric shadow.

An opaque particle can be described, for example, by the limit $m \to \infty$, i.e., a perfect conductor. One finds from Eq. (7.3) that in this case $\rho \to \infty$ regardless of kR and hence, from Fig. (11.5), $Q^{ext} \to 2$ implying that the size of an opaque particle need not be geometrically large for the paradox to occur. If one now considers the fact that the total wave behind an opaque particle of small to intermediate size is not necessarily zero, then it is again unclear why it should be necessary to identify a SFW to explain the paradox. Moreover, the generalization of the SFW concept to dielectric particles, does not show a physical explanation for *why* a SFW is needed for such semi-opaque particles [104].

[5]For the fractal aggregate, the ambiguity is due to the *smallness* of the constituent particles and their corresponding violation of geometrical optics.

[6]This appears to be especially true in quantum mechanical discussions of the paradox.

11.4 A New View of Extinction and its Paradox

Using either the diffraction or shadow explanation, one is able to predict the extinction behavior of geometrically large particles of any shape. In this regard these explanations are useful. However, given the discussions of Sec. (11.3), one is left with the impression that these concepts do not constitute a complete and convincing explanation for the *physical cause* of the paradox. The following will present a new description of a particle's extinction cross section and in the process reveal a simple and fundamental connection between the paradox and the Ewald-Oseen extinction theorem.

Using the volume integral equations Eqs. (2.10) and (2.18), the scattered field are

$$\mathbf{E}^{sca}(\mathbf{r}) = k^2(m^2-1)\int_{V^{int}} \overleftrightarrow{G}_e(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}^{int}(\mathbf{r}')\mathrm{d}\mathbf{r}' \quad \mathbf{r}\in V^{ext}, \tag{11.2}$$

$$\mathbf{B}^{sca}(\mathbf{r}) = \frac{-i\omega}{c}(m^2-1)\int_{V^{int}} \overleftrightarrow{G}_m(\mathbf{r},\mathbf{r}')\cdot\mathbf{E}^{int}(\mathbf{r}')\mathrm{d}\mathbf{r}' \quad \mathbf{r}\in V^{ext}. \tag{11.3}$$

The fields given by Eqs. (11.2) and (11.3) are exact expressions for the scattered wave for any location outside the particle. The volume integral in these expressions can be formally transferred to an integral over the particle surface $\mathcal{S}$ using Green's second theorem in vector-dyadic form, Eq. (9.13) see [90, Ch. 2], [30, p. 300]. A physical interpretation of the related vector Green's theorem is given in [106, Sec. 5.2]. The resulting surface integral takes on two values depending on whether $\mathbf{r}$ is inside or outside of the particle,

$$\oint_{\mathcal{S}} \left\{ i\omega\overleftrightarrow{G}_e(\mathbf{r},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{B}^{int}(\mathbf{r}')\right] + \overleftrightarrow{G}_m(\mathbf{r},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{E}^{int}(\mathbf{r}')\right]\right\}\mathrm{d}S' = \begin{cases} -\mathbf{E}^{inc}(\mathbf{r}) & \mathbf{r}\in V^{int} \\ \mathbf{E}^{sca}(\mathbf{r}) & \mathbf{r}\in V^{ext}. \end{cases} \tag{11.4}$$

In Eq. (11.4), $\hat{\mathbf{n}}$ is the outward surface normal to $\mathcal{S}$ and $\mathbf{r}\notin\mathcal{S}$. When $\mathbf{r}$ resides outside the particle, Eqs. (11.2) and (11.4) yield exactly the same scattered field. If however, $\mathbf{r}$ resides inside of the particle, Eq. (11.4) yields the negative of the incident field. This is one presentation of the Ewald-Oseen (EO) extinction theorem [90].

With Eqs. (11.2) and (11.4), one can now find equivalent and exact expressions for C^{ext} via Eq. (3.5) over its surface. Alternatively, one can take the observation point to reside in

the particle's far-field zone. This permits use of the far-field approximation to simplify the Green's functions in Eqs. (2.3) and (2.13) giving

$$\overleftrightarrow{G}_e(\mathbf{r},\mathbf{r}') = \frac{e^{ikr}}{r}\, g_e(\mathbf{r},\mathbf{r}') \quad kr \to \infty, \tag{11.5}$$

$$\overleftrightarrow{G}_m(\mathbf{r},\mathbf{r}') = \frac{e^{ikr}}{r}\, g_m(\mathbf{r},\mathbf{r}') \quad kr \to \infty, \tag{11.6}$$

where

$$\overleftrightarrow{g}_e(\hat{\mathbf{r}},\mathbf{r}') = \frac{1}{4\pi}\left(\overleftrightarrow{I} - \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\exp(-ik\mathbf{r}'\cdot\hat{\mathbf{r}}), \tag{11.7}$$

$$\overleftrightarrow{g}_m(\hat{\mathbf{r}},\mathbf{r}') = \frac{ik}{4\pi}\left(\hat{\mathbf{r}}\times\overleftrightarrow{I}\right)\exp(-ik\mathbf{r}'\cdot\hat{\mathbf{r}}). \tag{11.8}$$

By substituting Eqs. (11.5)-(11.8) in Eq. (11.4) and accounting for the factor $\exp(ikr)/r$, one can identify a far-field scattering amplitude in terms of an integral over the particle's surface

$$\mathbf{E}_1^{sca}(\hat{\mathbf{r}}) = \oint\limits_{\mathcal{S}} \left\{ i\omega \overleftrightarrow{g}_e(\mathbf{r},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{B}^{int}(\mathbf{r}')\right] + \overleftrightarrow{g}_m(\mathbf{r},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{E}^{int}(\mathbf{r}')\right]\right\} \mathrm{d}S'. \tag{11.9}$$

Now, using the optical theorem

$$C^{ext} = \frac{4\pi}{k|\mathbf{E}_o^{inc}|^2}\mathrm{Im}\left\{\left[\mathbf{E}_o^{inc}\right]^*\cdot\mathbf{E}_1^{sca}(\hat{\mathbf{n}}^{inc})\right\}, \tag{11.10}$$

the extinction cross section is

$$C^{ext} = \frac{4\pi}{k|\mathbf{E}_o^{inc}|^2}\mathrm{Im}\oint\limits_{\mathcal{S}} \left[\mathbf{E}_o^{inc}\right]^*\cdot\left\{ i\omega \overleftrightarrow{g}_e(\hat{\mathbf{n}}^{inc},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{B}^{int}(\mathbf{r}')\right] + \overleftrightarrow{g}_m(\hat{\mathbf{n}}^{inc},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{E}^{int}(\mathbf{r}')\right]\right\} \mathrm{d}S'. \tag{11.11}$$

From Eq. (11.1), Q^{ext} can be found by dividing Eq. (11.11) by G.

The surface integral in Eq. (11.11) provides the opportunity to see how different regions of the particle surface $\mathcal{S}$ contribute to C^{ext} and Q^{ext}. To do this, the integral is evaluated only over the portion of the particle's surface extending from $\theta = \pi$ to $\theta = \theta_s$. This partial particle-surface, denoted $\partial\mathcal{S}$, is shown in Fig. (10.4) and is reminiscent of $\partial\mathcal{S}_b$ in Fig. (10.4),

except here the surface resides on the particle. Evaluating Eq. (11.11) over $\partial\mathcal{S}$ and using Eq. (11.1) allows one to define a partial extinction efficiency factor ∂Q^{ext} as

$$\partial Q^{ext}(\theta_s) = \frac{1}{G}\frac{4\pi}{k|\mathbf{E}_o^{inc}|^2}\mathrm{Im}\int_{\partial\mathcal{S}}\left[\mathbf{E}_o^{inc}\right]^*\cdot\Big\{i\omega\overleftrightarrow{g}_e(\hat{\mathbf{n}}^{inc},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{B}^{int}(\mathbf{r}')\right] + \overleftrightarrow{g}_m(\hat{\mathbf{n}}^{inc},\mathbf{r}')\cdot\left[\hat{\mathbf{n}}\times\mathbf{E}^{int}(\mathbf{r}')\right]\Big\}\mathrm{d}S'. \quad (11.12)$$

When $\theta_s = \pi$, there is no surface and the integral in Eq. (11.12) is zero, whereas when $\theta_s = 0$ the partial surface becomes the complete particle surface $\partial\mathcal{S} = \mathcal{S}$ and Eq. (11.12) yields Q^{ext}. Note that θ_s closes the surface as it decreases from π to zero.

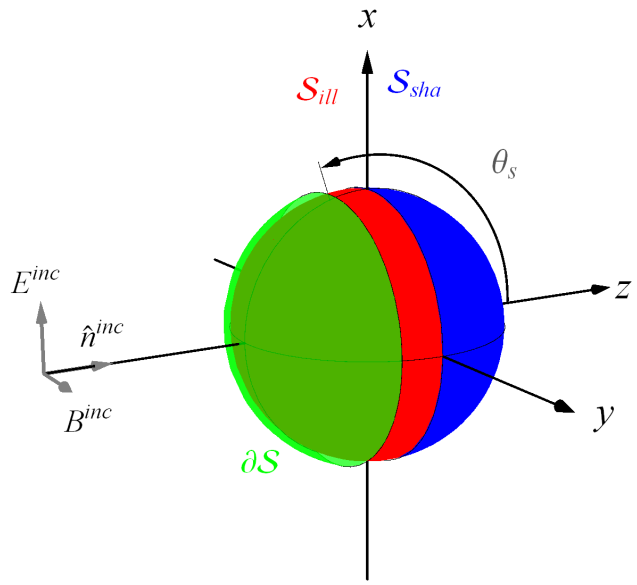

Figure 11.7: Partial particle-surface $\partial\mathcal{S}$ used to calculate ∂Q^{ext} in Eq. (11.12) and ∂G in Eq. (11.13). The geometrically illuminated and shaded sides of the particle, $\mathcal{S}_{ill}$ and $\mathcal{S}_{sha}$, are shown in red and blue, respectively.

Figure (11.8) shows ∂Q^{ext} as a function of decreasing θ_s for spherical particles of various kR and m, which are chosen so that each particle belongs to one of three values of ρ representative of the Rayleigh, resonance, and asymptotic regions of Q^{ext}, recall Sec. (11.1), Fig. (11.5), and Eq. (7.3). Also shown with the ∂Q^{ext} curves is the particle's partial

geometrical projection ∂G

$$\partial G(\theta_s) = -\int_{\partial\mathcal{S}} \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^{inc}\, \mathrm{d}S, \tag{11.13}$$

which gives the projection of $\partial\mathcal{S}$ into the forward direction. When $\theta_s = \pi/2$, this projection is equivalent to G. Each particle's Q^{ext} as calculated directly from the Mie series is indicated in Fig. (11.8). Comparison between these values of Q^{ext} and those resulting from Eq. (11.12) verifies that Eq. (11.12) does indeed yield the expected values for Q^{ext} as $\theta_s \to 0$.

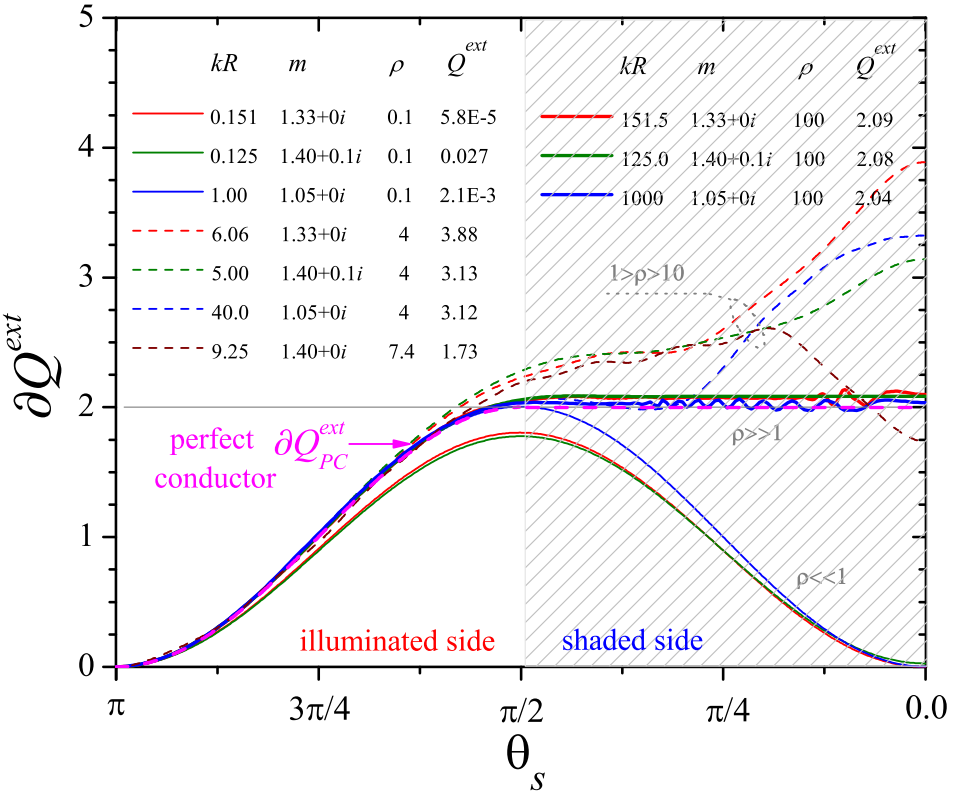

Figure 11.8: Plots of ∂Q^{ext} for various spherical particles as a function of θ_s, recall Fig. (11.7). The relative size kR, refractive index m, phase shift ρ, and value of Q^{ext} (as calculated from Mie theory) are shown for each particle in the legend. Also show from $\theta_s = \pi$ to $\theta_s = \pi/2$ is the partial particle-shape-projection ∂G of Eq. (11.13).

One can see in Fig. (11.8) that the curves rise from zero at $\theta_s = \pi$ to around unity when $\theta_s = \pi/2$. Moreover, all the curves roughly follow the partial geometrical projection ∂G in this range. As θ_s decreases past $\pi/2$ to $\theta_s = 0$, the curves separate into three groups. For $\rho < 1$, the curves fall back to a small value for Q^{ext} at $\theta_s = 0$. If $\rho \gtrsim 1$, the curves appear

uncorrelated and spread out over a large range of values for Q^{ext}. If however $\rho \gg 1$, the curves bundle together and roughly hold a constant value of $\partial Q^{ext} \sim 2$ as $\pi/2 > \theta_s \geqslant 0$.

To further investigate Fig. (11.8) one can look at the differential contributions to Q^{ext} from the surface integral in Eq. (11.12). This is done by discretizing $\mathcal{S}$ into small surface elements $\Delta\mathcal{S}$ and identifying a complex number, or *surface phasor*, that represents a surface element's contribution to the integral. Equation (11.12) is then

$$Q^{ext} = \frac{1}{G}\frac{4\pi}{k|\mathbf{E}_o^{inc}|^2}\mathrm{Im}\Big\{\sum_i z_i^{(ill)}\Delta\mathcal{S}_i + \sum_j z_j^{(sha)}\Delta\mathcal{S}_j\Big\} \tag{11.14}$$

where $z_i^{(ill)}$ and $z_i^{(sha)}$ are the surface phasors corresponding to the illuminated and shaded hemispheres $\mathcal{S}_{ill}$ and $\mathcal{S}_{sha}$, respectively, and are given by

$$\lim_{\Delta\mathcal{S}\to 0}\sum_i z_i^{\binom{ill}{sha}}\Delta\mathcal{S}_i = \int_{\mathcal{S}_{\binom{ill}{sha}}} \big[\mathbf{E}_o^{inc}\big]^* \cdot \Big\{i\omega\overleftrightarrow{g}_e(\hat{\mathbf{n}}^{inc},\mathbf{r}')\cdot\big[\hat{\mathbf{n}}\times\mathbf{B}^{int}(\mathbf{r}')\big] + \overleftrightarrow{g}_m(\hat{\mathbf{n}}^{inc},\mathbf{r}')\cdot\big[\hat{\mathbf{n}}\times\mathbf{E}^{int}(\mathbf{r}')\big]\Big\}\mathrm{d}S'. \tag{11.15}$$

The phasors $z_i^{(ill)}$ and $z_i^{(sha)}$ in Eq. (11.14) can be given a clear meaning. By making an analogy to the interpretation of the VIE described in Ch. 4, one can regard the tangential field components $\hat{\mathbf{n}}\times\mathbf{B}^{int}(\mathbf{r}')$ and $\hat{\mathbf{n}}\times\mathbf{E}^{int}(\mathbf{r}')$ in Eq. (11.11), (11.12), and (11.15) as "effective surface sources" that radiate a spherical wave to the far-field under the action of the functions $\overleftrightarrow{g}_e$ and $\overleftrightarrow{g}_m$. These sources are "effective" because they are not real wavelet sources; they are ultimately a mathematical consequence of the transformation of the volume integral in Eqs. (11.2) and (11.3) resulting from application of Green's theorem, Eq. (9.13), leading to Eq. (11.4). These surface sources, however, are determined by the real wavelet sources that reside through the particle volume. Consequently, one can think of the collection of $z_i^{(ill)}$ and $z_i^{(sha)}$, corresponding to the entire particle surface, as an equivalent representation of the actual wavelet sources inside the particle. In addition, by comparing the form of the complex exponentials in Eqs. (3.1), (11.7), and (11.8) and noting that it is the complex conjugate of the incident electric field amplitude that appears in Eq. (11.15), one can see

that the phase angle of the phasors $z_i^{(ill)}$ and $z_i^{(sha)}$ represents the *relative* phase between the effective surface sources and the incident field at these source's location on the particle surface.

Plots (a) and (b) in Figures (11.9)-(11.11) show $z_i^{(ill)}$ and $z_i^{(sha)}$ in the complex plane for spherical particles with $m = 1.33 + 0i$ and sizes $kR = 0.151$, $kR = 6.06$, and $kR = 151$, respectively. The plots labeled (c) and (d) in these figures show the distribution of the phase *angles* of $z_i^{(ill)}$ and $z_i^{(sha)}$ over $\mathcal{S}_{ill}$ and $\mathcal{S}_{sha}$. The magnitude of the phasors in (a) and (b) are collectively normalized to unity and the phase angles in plots (c) and (d) are indicated by the color scale shown in the legend. The same size and refractive index particles in Figs. (11.9)-(11.11) also appear in Fig. (11.8).

Consider first the curve for ∂Q^{ext} in Fig. (11.8) for the Rayleigh-sized sphere with $kR = 0.151$ and $m = 1.33 + 0i$, and compare to the surface-phasor plots for this same particle in Fig. (11.9). One can see that the relative phase over the particle's illuminated side is roughly $\pi/2$, whereas it is roughly $-\pi/2$ over the shaded side. Now consider the ∂Q^{ext} curve in Fig. (11.8) for the resonance-sized sphere with $kR = 6.06$ and compare to the corresponding phasor plots in Fig. (11.10). Here one sees that, as with the Rayleigh-sized particle, the relative phase over the illuminated side is roughly constant and equal to $\pi/2$. However, now the relative phase over the shaded side is no longer roughly constant and varies over the complex plane. Finally, consider the ∂Q^{ext} curve in Fig. (11.8) for the asymptotic-sized sphere with $kR = 151$ and compare to its corresponding phasor plots in Fig. (11.11). Here, the relative phase is again seen to be roughly constant and equal to $\pi/2$ over the illuminated hemisphere $\mathcal{S}_{ill}$. Now though, the phase over the shaded hemisphere varies extensively and covers the entire complex plane in a roughly isotropic distribution.

It is now possible to make a general conclusion with regard to the asymptotic behavior of Q^{ext} based on Figs. (11.8)-(11.11); if $\rho \gg 1$, *the geometrically illuminated region of a particle makes a contribution of roughly* $2G$ *to* C^{ext} *while its geometrically shaded region makes a comparatively negligible contribution.* Moreover, one can understand from Fig.

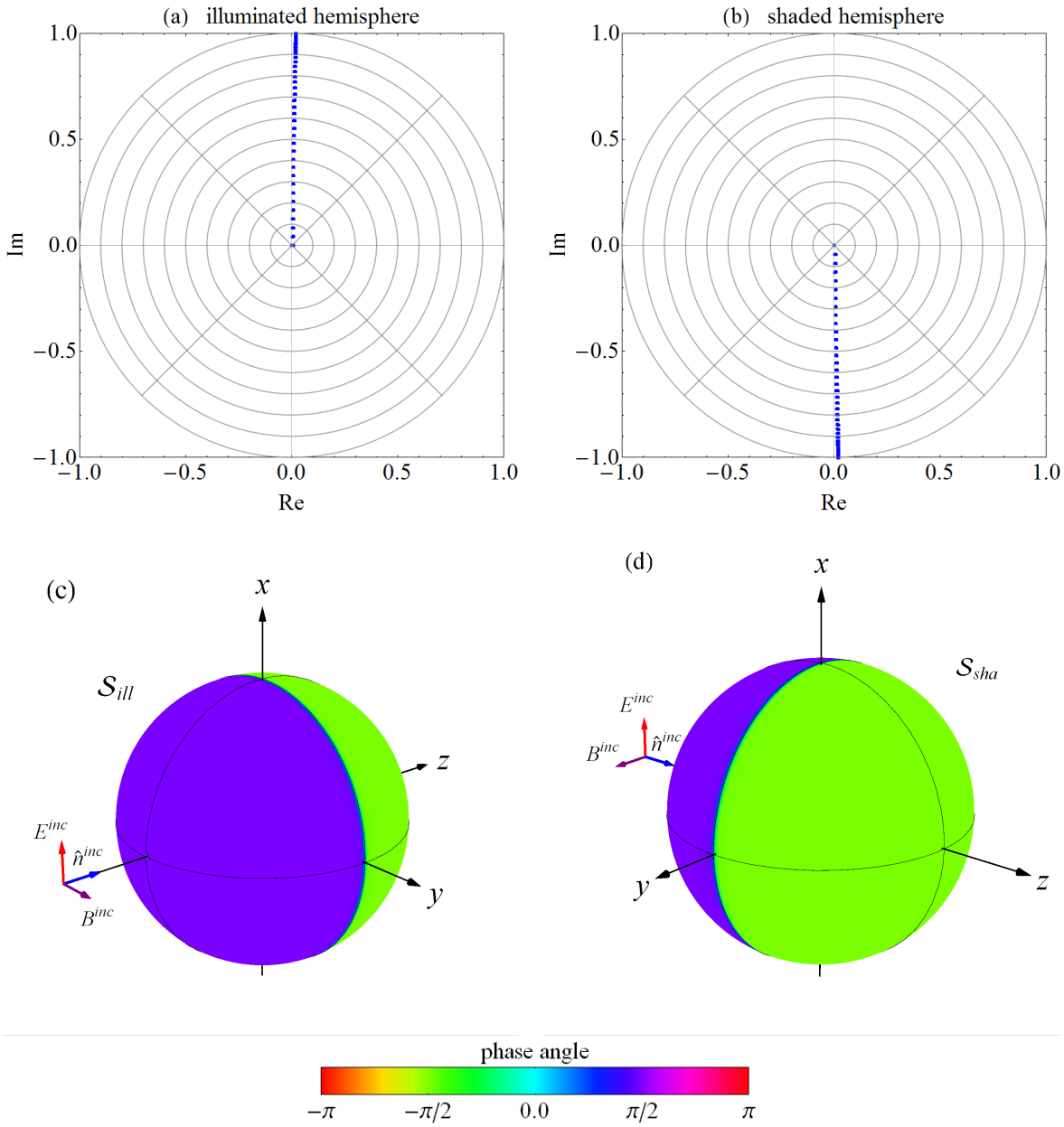

Figure 11.9: Plots (a) and (b) show the phasors of Eq. (11.14) in the complex plane for the illuminated and shaded surfaces of the particle. Plots (c) and (d) show the distribution of phase angle of these phasors over the particle's surface as indicated by the color scale shown in the legend. The particle in these plots has $kR = 0.151$ and $m = 1.33 + 0i$ and hence is a member of the Rayleigh-region in Fig. (11.8). The phasors are globally scaled to a maximum magnitude near one for clarity.

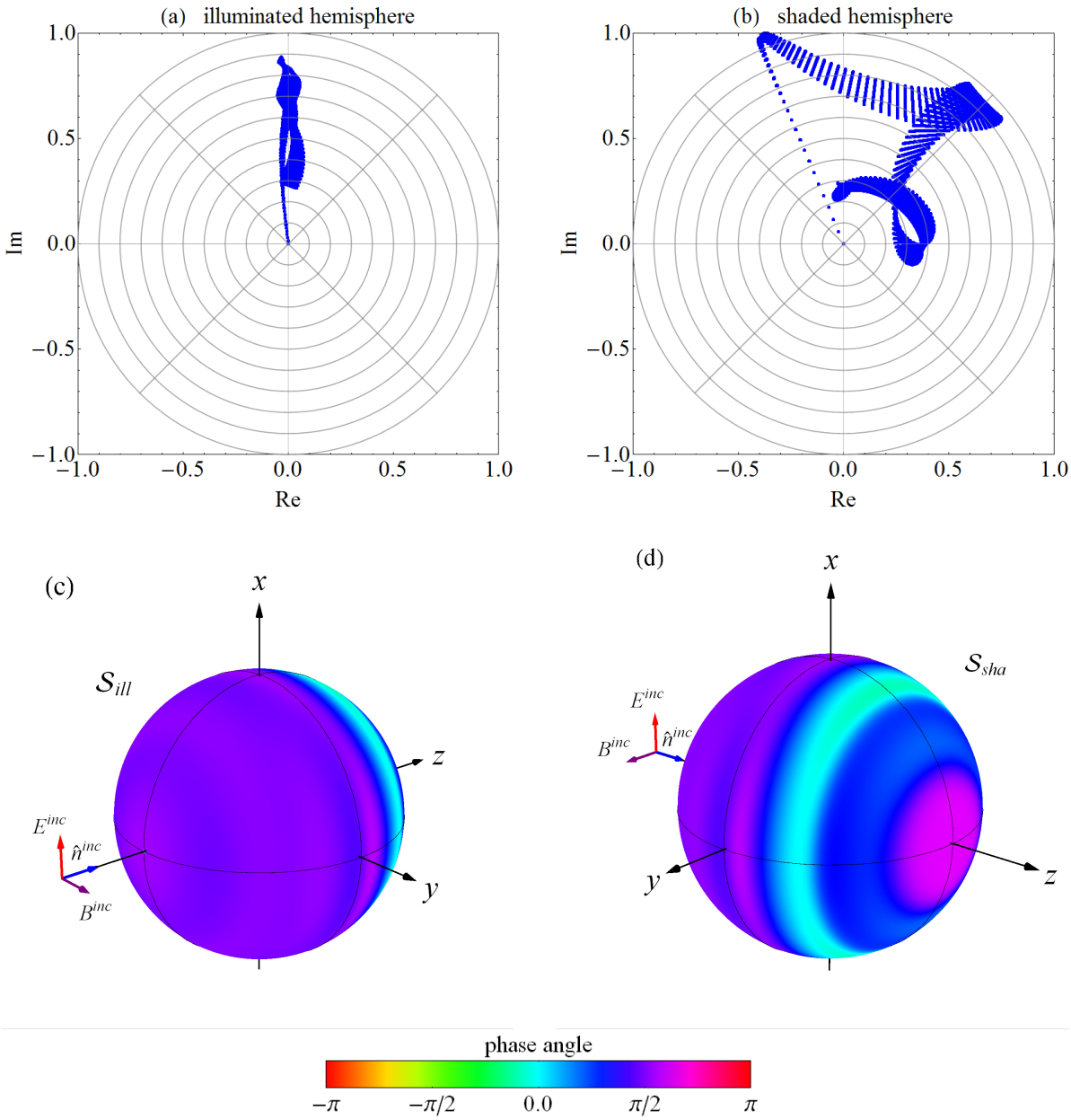

Figure 11.10: Same as Fig. (11.9) except the particle is now representative of the resonance-region of Fig. (11.8), with $kR = 6.06$ and $m = 1.33 + 0i$.

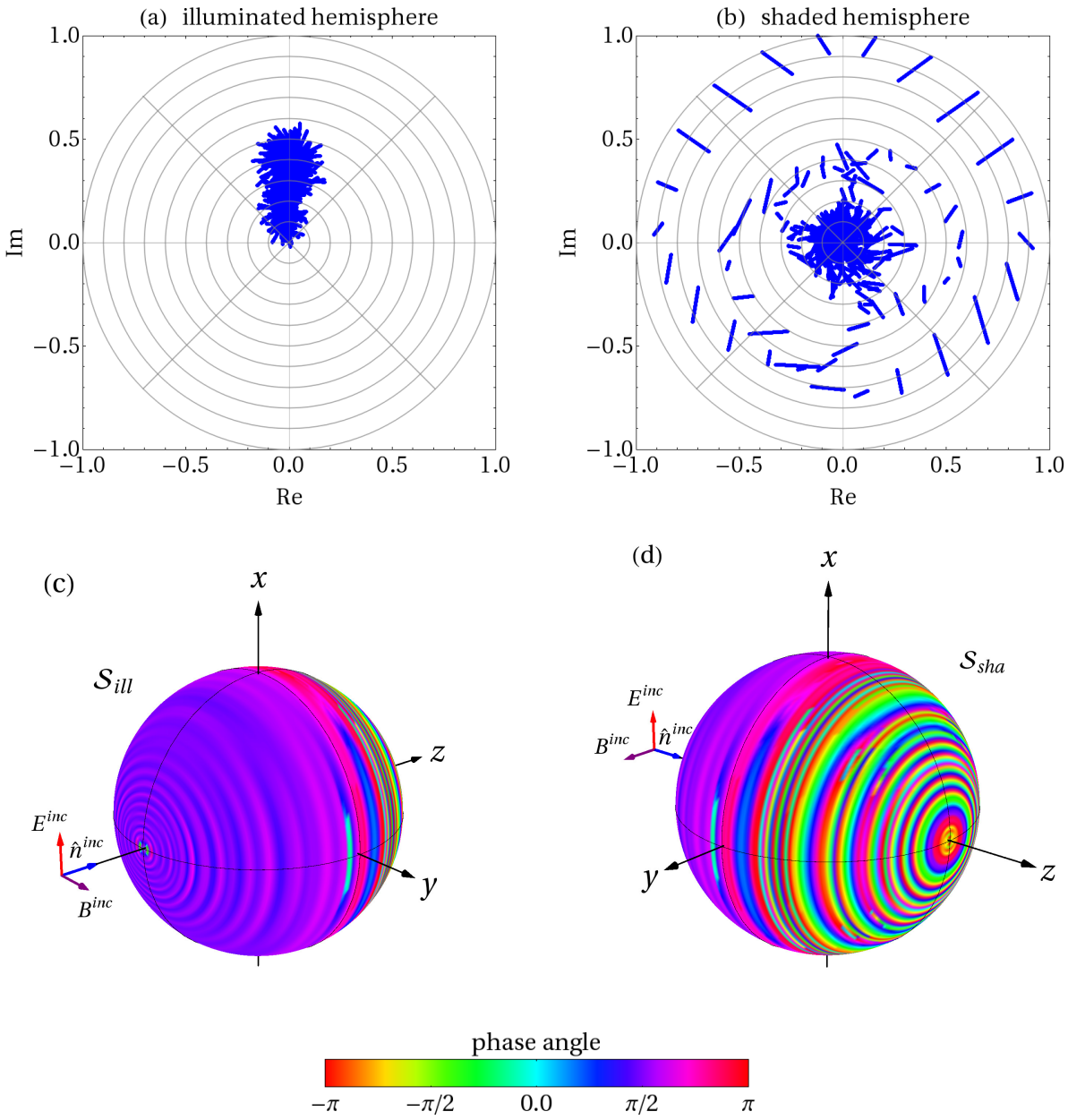

Figure 11.11: Same as Figs. (11.9) and (11.10) except this particle belongs to the asymptotic-region in Fig. (11.8) with $kR = 151$ and $m = 1.33 + 0i$.

(11.11) that the reason the shaded region makes no significant contribution is due to the near-complete cancellation of the $z_i^{(sha)}$ phasors in the complex plane, i.e., near-complete destructive interference. Notice also, from Fig. (11.8), that this conclusion applies to any kR and m so long as $\rho \gg 1$; *it is not necessarily restricted to geometrically large or opaque particles.*

The final step required to reveal the full significance of the extinction paradox is to combine the understanding provided by Figs. (11.8)-(11.11) with the concept of the Ewald-Oseen (EO) extinction theorem. The embodiment of the EO theorem is given by Eq. (11.4), which shows that the incident field is canceled within V^{int} by the surface integral on the left hand side of that equation [90, p. 62-63], [107, Ch. 12]. This same surface integral appears in its far-field form in Eq. (11.11) yielding C^{ext}. Consequently, the conclusions predicated on Eq. (11.11) and Figs. (11.8)-(11.11) suggest a connection between EO extinction and the dominance of the particle's contribution to C^{ext} originating from its geometrically illuminated side.

To further examine the connection between the EO theorem and C^{ext}, Figs. (11.12)-(11.14) show plots of the electric field magnitude in the x-z plane[7] through the center of the same three spherical particles as in Figs (11.9)-(11.11). The incident electric field is polarized parallel to the x-axis and propagates along the z-axis in the positive direction in these figures, recall Fig. (1.1). Each figure shows four plots with the outline of the particle's surface indicated by the dashed white line. Plot (a) displays the *total* electric field magnitude outside and inside the particle. Plot (b) shows the magnitude of the superposition of the incident field and the field generated by the surface integral in Eq. (11.4) when the integral is evaluated over $\mathcal{S}_{ill}$ only. The outline of $\mathcal{S}_{ill}$ is shown in red in the plot. Similarly, plot (c) shows the magnitude of the superposition of the incident field and the surface-integral field, except here the integral is evaluated over $\mathcal{S}_{sha}$ only. The outline of $\mathcal{S}_{sha}$ is shown in blue. Lastly, plot (d) shows the same superposition except now the integral is evaluated over the

[7] This is the same as the vertical scattering plane Π_y in Fig. (6.2).

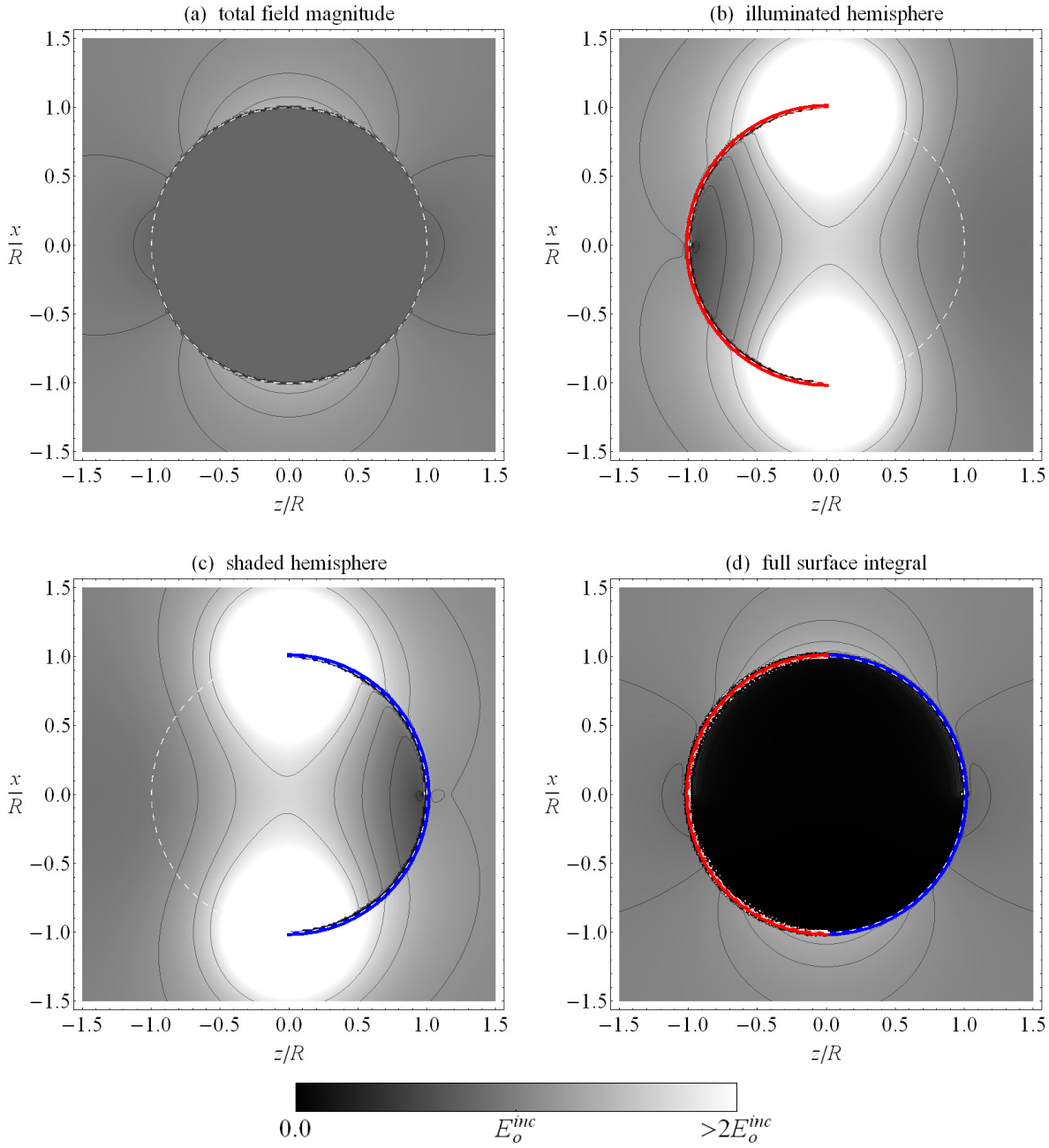

Figure 11.12: Plot (a) shows the electric field magnitude in the x-z plane passing through the center of a spherical particle with $kR = 0.151$ and $m = 1.33 + 0i$. Plots (b) and (c) show the field magnitude as calculated from Eq. (11.4) when the surface integral in that equation is evaluated over only $\mathcal{S}_{ill}$ or $\mathcal{S}_{sha}$, respectively. The intersection of these surfaces with the x-z plane is shown in the plots in red for the illuminated side $\mathcal{S}_{ill}$, and blue of the shaded side $\mathcal{S}_{sha}$. Plot (d) shows the field magnitude produced by the full surface integral. The gray scale in the legend indicates the field magnitude relative to that of the incident wave E_o^{inc}.

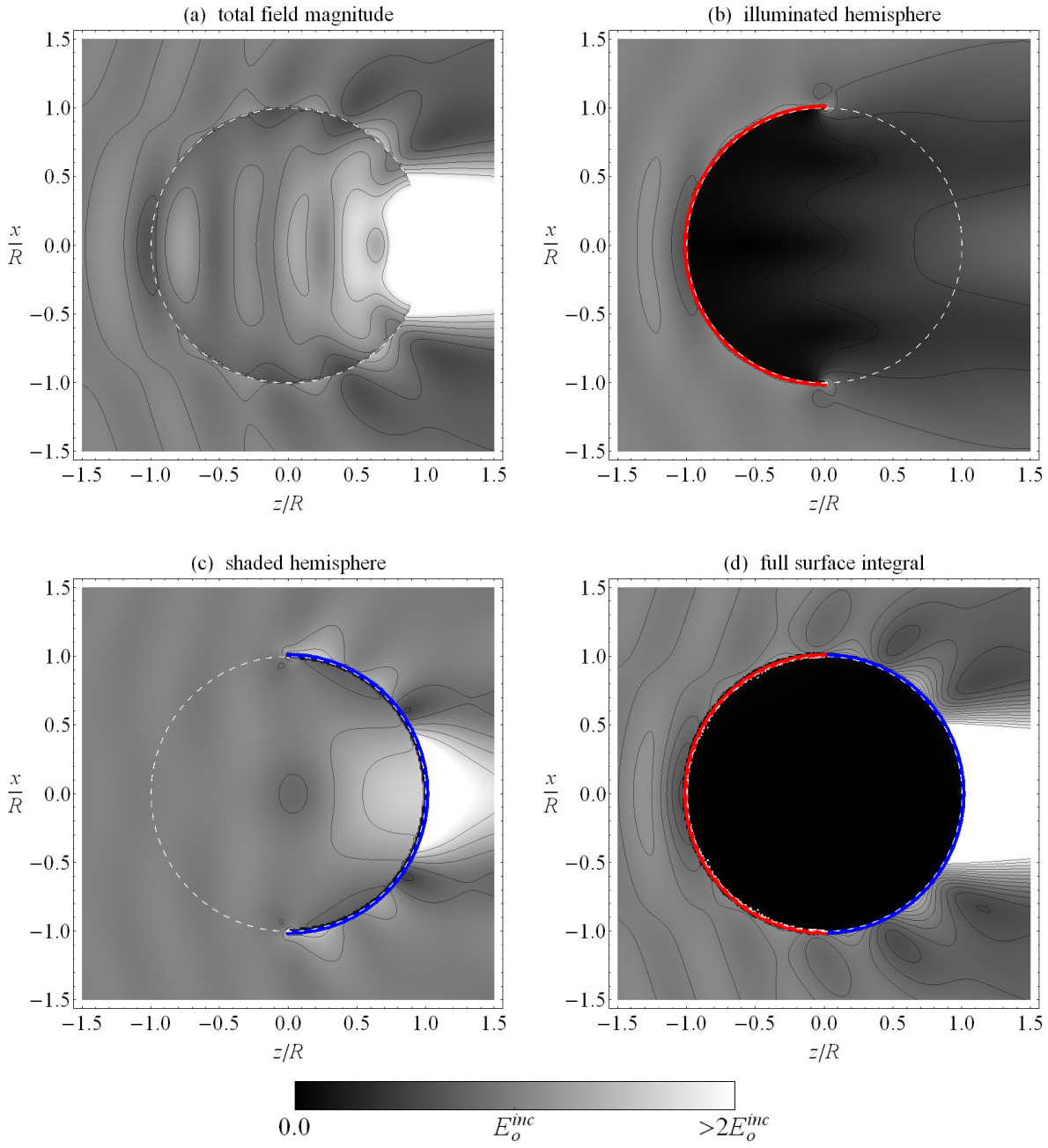

Figure 11.13: Same as Fig. (11.12) except for a particle with $kR = 6.06$ and $m = 1.33 + 0i$ corresponding to the resonance-region of Fig. (11.8).

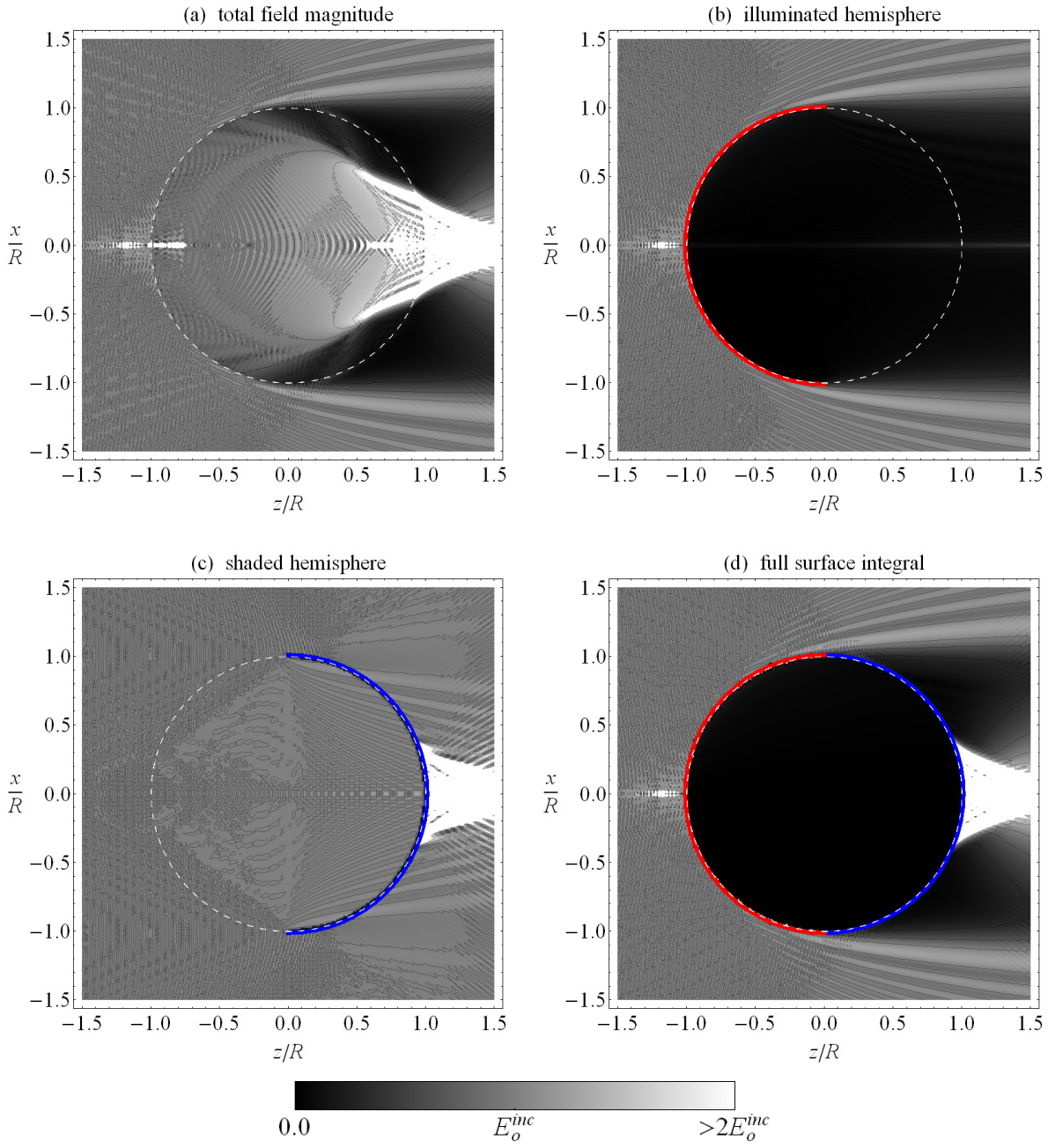

Figure 11.14: Same as Figs. (11.12) and (11.13) except for a particle with $kR = 151$ and $m = 1.33 + 0i$ corresponding to the asymptotic-region of Fig. (11.8).

complete particle surface $\mathcal{S}$.

The sphere shown in Fig. (11.12) is the same sphere as in Figs. (11.8) and (11.9) corresponding to the Rayleigh-region with $kR = 0.151$ and $m = 1.33 + 0i$. From plot (a) in Fig. (11.12) one can see the familiar electric-dipole-like field distribution in V^{ext} and a roughly constant field magnitude in V^{int}. Plot (d) shows that the surface integral in Eq. (11.4) does indeed cancel out the incident field inside the particle and produces the same external field outside. *This cancellation is a graphical demonstration of the Ewald-Oseen extinction theorem.* Plots (b) and (c), however, clearly show that evaluation of this surface integral over either $\mathcal{S}_{ill}$ or $\mathcal{S}_{sha}$ is incapable of canceling the incident field in V^{int} except over a small portion of V^{int} near the $z = R$ and $z = -R$ ends of the $\mathcal{S}_{ill}$ or $\mathcal{S}_{sha}$ regions. Neither plot (b) or (c) are capable of generating the expected total wave in V^{ext} either.

Now consider Fig. (11.13) showing a larger sphere with $kR = 6.06$ and $m = 1.33 + 0i$, which is the same sphere as in Figs. (11.8) and (11.10). Here again the full surface integral, (d), is seen to cancel the incident field in V^{int} and produce the same total field as seen in V^{ext} in (a). Neither half-surface integral in (b) or (c) generates a correct total field outside of the particle, but in (b) the incident field is almost completely canceled within the portion of the particle interior encompassed by the $\mathcal{S}_{ill}$ cap. Incident-field cancellation is not seen in (c).

Lastly, consider the asymptotic-region sphere of Fig. (11.14) with $kR = 151$ and $m = 1.33 + 0i$. This is the same sphere as in Fig. (11.11) and is included in Fig. (11.8). Once again, one can see the agreement between the total field in V^{ext} in (a) and (d) and the cancellation of the incident field in V^{int} in (d). *However, notice in (b) that the surface integral evaluated over only $\mathcal{S}_{ill}$ is capable of canceling the incident field in all V^{int} just as the full surface integral in (d) does.* One can see from (c) that the integral over $\mathcal{S}_{sha}$ has almost no ability to cancel the incident field in V^{int} but dominates the total field in V^{ext} expected from (a), especially in the forward direction.

11.5 Interpretation

From the observations in Sec. (11.4), it is now possible to develop a new understanding for the cause of the extinction paradox. The curves corresponding to the asymptotic region in Fig. (11.8) combined with Fig. (11.11) demonstrate that the geometrically illuminated region of a particle is primarily responsible for its extinction cross section if $\rho \gg 1$. Moreover, Fig. (11.14) shows that this same region is primarily responsible for achieving EO cancellation of the incident wave. Therefore, one might expect that there must be a connection between the requirement that the incident wave be canceled in V^{int} and the result that $C^{ext} \to 2G$ as $\rho \to \infty$. This connection will be described below, and in the process one will see how the EO theorem explains the extinction paradox.

A convenient way to attain the $\rho \to \infty$ limit is to consider a spherical particle composed of a perfect electric conductor (PEC). In this case, one knows that the internal fields are zero and that a current $\mathbf{K}$ will be induced on the particle's surface corresponding to the discontinuity in the tangential component of the magnetic field, Eq. (1.40),

$$\hat{\mathbf{n}} \times \mathbf{B}^{inc}(\mathbf{r}) + \hat{\mathbf{n}} \times \mathbf{B}^{sca}(\mathbf{r}) = \mu_o \mathbf{K}(\mathbf{r}) \quad \mathbf{r} \in \mathcal{S}.$$

One can find the field radiated by this surface current using the three-dimensional analog to Eq. (9.43),

$$i\omega\mu_o \oint_{\mathcal{S}} \overleftrightarrow{G}_e(\mathbf{r}, \mathbf{r}') \cdot \mathbf{K}(\mathbf{r}') \mathrm{d}S' = \begin{cases} -\mathbf{E}^{inc}(\mathbf{r}) & \mathbf{r} \in V^{int} \\ \mathbf{E}^{sca}(\mathbf{r}) & \mathbf{r} \in V^{ext} \end{cases}, \tag{11.16}$$

see also [30, p. 64]. In general, current is induced over the entire particle surface. However, if the particle is large, the current over the shaded hemisphere $\mathcal{S}_{sha}$ tends to zero [108, p. 205-213]. The physical optics approximation is then introduced, which assumes that the current is zero in the shaded region [109]. In addition, the large size of the particle allows one to approximate each differential surface element as an infinite plane, in which case [15, Sec 10.10]

$$\hat{\mathbf{n}} \times \mathbf{B}^{sca}(\mathbf{r}) = \hat{\mathbf{n}} \times \mathbf{B}^{inc}(\mathbf{r}) \quad \mathbf{r} \in \mathcal{S} \quad kR \to \infty. \tag{11.17}$$

Using Eqs. (3.1), (3.2), and (11.17) in combination with Eq. (1.40), gives

$$\mathbf{K}(\mathbf{r}) = \begin{cases} 2\frac{k}{\omega\mu_o}\hat{\mathbf{n}} \times \left[\hat{\mathbf{n}}^{inc} \times \mathbf{E}^{inc}(\mathbf{r})\right] & \mathbf{r} \in \mathcal{S}_{ill} \\ 0, & \mathbf{r} \in \mathcal{S}_{sha}. \end{cases} \tag{11.18}$$

Combining Eqs. (11.16) and (11.18), taking the far-field limit of the Green's function, and using the optical theorem Eq. (11.10), one eventually finds that

$$C^{ext} = 2\mathrm{Im}\left\{ -i \int_{\mathcal{S}_{ill}} \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^{inc} \mathrm{d}S \right\} = 2G, \tag{11.19}$$

which is the expected result.

The calculation leading to Eq. (11.19) is simple enough that it is possible to see the significance of the factor 2. The first appearance of this factor is in the surface current of Eq. (11.18). This form for the current is a consequence of the boundary conditions and the implicit requirement that the wave inside the particle be zero. Combination of Eqs. (1.40) and (11.18) along with the physical optics approximation show that the tangential component of the total magnetic field on $\mathcal{S}_{ill}$ is *twice* the corresponding magnitude of the same field component that would exist in the particle's absence. *The induced surface source* ($\mathbf{K}$) *must be twice the magnitude expected the incident wave only such that it be able to cancel the wave inside the particle and concomitantly produce a reflected wave of similar magnitude.* The magnitude of the tangential component of the surface field must be twice that of the incident field such that the induced current is able to radiate a wave into the particle that cancels the incident and simultaneously radiates a scattered wave of similar magnitude outside. Note the striking similarity of this realization to the discussions in Secs. 9.1 and 9.6. Since the cross section is ultimately determined by the induced current, this result explains the factor 2 in C^{ext}.

Notice that there are clear similarities here to the dielectric particles considered in Sec (11.4). Both a dielectric particle with $\rho \gg 1$ and a perfectly conducting particle involve only their geometrically illuminated surfaces $\mathcal{S}_{ill}$ in C^{ext}. With regard to its contribution to C^{ext}, the destructive interference of the differential contributions from the shaded region in a dielectric particle and the corresponding zero-magnitude contribution from a perfectly

conducting particle are equivalent. Recall Fig. (11.11) and Eq. (11.18). Most importantly, *the requirement that the field be zero inside the perfectly conducting particle is inclusive of the requirement that the incident field be canceled there*; this finally provides a connection between the EO theorem and the asymptotic behavior of C^{ext} with ρ.

The Ewald-Oseen theorem imposes a special requirement of the internal field; the wavelet sources that constitute it, must be induced and couple together such that the secondary radiation from them is able to cancel the incident wave throughout the particle interior[8]. This property of the wavelet-sources is then formally connected to C^{ext} through Eqs. (3.5), (11.2), and (11.3). In the $\rho \to \infty$ limit, the incident-wave cancellation is achieved by the particle's geometrically illuminated region, hence explaining why the paradox depends only on G. Moreover, in analogy to the perfectly conducting particle, EO cancellation requires that the particle's *effective* surface wavelet-sources $\hat{\mathbf{n}} \times \mathbf{E}^{int}$ and $\hat{\mathbf{n}} \times \mathbf{B}^{int}$ have twice the magnitude as would be expected in the particle's absence. Since C^{ext} is directly related to the magnitude of these wavelet sources [Eq. (11.11)], this explains why $C^{ext} = 2G$ and not $C^{ext} = G$. *It is the EO extinction theorem that accounts for the unexpected extra factor of G in the extinction paradox.*

Inspection of the curves in Fig. (11.8) for $\rho \gg 1$ over the range $\pi \geqslant \theta_s > \pi/2$ shows a strong functional similarity to the projection integral of Eq. (11.13) i.e.,

$$C^{ext} = \frac{4\pi}{k|\mathbf{E}_o^{inc}|^2}\mathrm{Im}\oint_{\mathcal{S}_{ill}} \left[\mathbf{E}_o^{inc}\right]^* \cdot \Big\{ i\omega \overleftrightarrow{g}_e(\hat{\mathbf{n}}^{inc}, \mathbf{r}') \cdot \left[\hat{\mathbf{n}} \times \mathbf{B}^{int}(\mathbf{r}')\right] + \overleftrightarrow{g}_m(\hat{\mathbf{n}}^{inc}, \mathbf{r}') \cdot \left[\hat{\mathbf{n}} \times \mathbf{E}^{int}(\mathbf{r}')\right] \Big\} \mathrm{d}S'$$
$$= 2\mathrm{Im}\Big\{ -i \int_{\mathcal{S}_{ill}} \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^{inc} \mathrm{d}S \Big\} = 2G \quad \rho \to \infty. \quad (11.20)$$

This result is remarkable because it applies to dielectric as well as opaque particles. From plot (a) in Fig. (11.14), one can see that the particle has large internal field magnitudes in both the illuminated and shaded regions and in the geometrical shadow. In contrast, a

[8]Note that this is why the incident wave is not present with the internal wave in the discussion in Sec. 5.6

perfectly conducting particle has no internal field and a near-field geometrical shadow region that would look much like plot (b) in Fig. (11.14). Yet, despite the striking differences between the internal fields in these types of particles, in the $\rho \to \infty$ limit, their cross section ultimately depends only on the special character of the internal fields that achieves EO cancellation of the incident wave in V^{int}. This interpretation is supported in Eq. (11.20) by the effective equivalence of the two surface integrals, and is illustrated by the similarities between plot (b) in Fig. (11.14) and what the corresponding plot (a) would look like for an opaque particle.

The reason that the geometrically shaded side of the particle makes no direct contribution to C^{ext} in the asymptotic region is because the largeness of ρ corresponds to strong refraction. This refraction causes the phase of the internal field, relative to the incident wave, to advance rapidly throughout the particle interior. This is why the phasors in Fig. (11.11) plot (b) are spread over the entire complex plane, corresponding to destructive interference between them. Although strong refraction occurs within the illuminated region as well, the EO theorem requires that the phase in this region advance relative to the incident wave in such a way as to achieve cancellation of the wave throughout V^{int}. This explains why the phasors in plot (a) of Fig. (11.11) are clustered along the positive imaginary axis and hence contribute to C^{ext} through the Im$\{\ldots\}$ filter in the optical theorem Eq. (11.10), see also [110, p. 406-10]. The same relative phase shift of $\pi/2$ is seen in the induced surface current on the infinite PC plate of Ch. 9; there to the incident wave is canceled because of this wavelet-source phase shift.

11.6 The "Flower Pot"

It is now of interest to consider the implications of the aboven EO-based explanation of the paradox in the context of a practical example. In describing the diffraction explanation, van de Hulst states [47, p. 107],

"A flower pot in a window prevents only the sunlight falling on it from entering

the room, and not twice this amount, but a meteorite of the same size somewhere in interstellar space between a star and one of our big telescopes will screen twice this light."

In this context, the EO explanation appears to predict that the flower pot should also "screen," i.e., extinct, twice the amount of light from entering the room. This is because, with regard to the EO theorem, there is no difference between the flower pot and the meteorite.

It is tacitly assumed in this example that either object's extinction cross section corresponds to the reduction of power received by a detector facing into the forward direction when the object is present as compared to the power received when the object is absent. This is the operational definition of extinction[9] and would imply that the flower pot have $C^{ext} = G$ as opposed to the expected $C^{ext} = 2G$. Consequently, van de Hulst adds that [47],

"...the observation [of C^{ext}] is made at very great distance, i.e., far beyond the zone where a shadow can be distinguished."

However, the general calculations in Ch. 3 prove that the cross section is independent of the distance from the particle and hence a *proper* measurement of the full cross section must still find $C^{ext} = 2G$ for the flower pot.

The analytical definition of extinction gives C^{ext} as the net energy flow, due to the interference between the incident and scattered waves, that passes through the surface $\mathcal{S}_{en}$ defining the system, recall Secs. 10.2, 10.3, and Eq. (3.5). Therefore, since the room in the flower pot example is not in the pot's far-field zone, *the reduced amount of light entering the room when the pot is in the window does not necessarily correspond to the pot's extinction cross section*, as van de Hulst says. Extinction has the character of being the forward-angle effect implied by the operational definition only in the far field, recall Ch. 10.

Although the room, which tacitly represents the detector, does not reside in the flower pot's far-field, it should nevertheless be possible to see how one can establish that $C^{ext} = 2G$

[9] See [35].

from a simple measurement. This is possible by recalling the energy conservation condition

$$C^{ext} = C^{sca} + C^{abs}. \tag{11.21}$$

Now suppose, for example, that the flower pot is made of metal and hence can be regarded as a perfectly conducting particle. This would correspond to $C^{abs} = 0$ since a perfect conductor absorbs no energy, and from Eq. (11.21) one sees that $C^{ext} = C^{sca}$. Therefore, if it is possible to account for the energy flow in the flower pot's *scattered wave*, constituting C^{sca}, C^{ext} will follow. First, notice that the amount of power contained in the scattered wave within the pot's geometrical shadow must be equal to that of the incident wave. This is because the scattered wave must cancel the incident wave there; *the observation of a well-defined shadow behind the pot indicates a contribution to* C^{sca} *equal to* G. This *is* the very concept of the shadow forming wave mentioned in Sec. (11.2.2). An additional factor of G is supplied to C^{sca} from reflection of the light geometrically incident on the pot's illuminated side, c.f. Fig. (11.4). Thus, the total scattering cross section is $C^{sca} = 2G$ giving $C^{ext} = 2G$ as expected. *There is no need to account for any concept of "diffraction" from the pot's shadow boundary.*

Appealing to the discussions in Sec. (11.4), one can now understand that the shadow forming wave in the above example is really due to EO cancellation *within* the pot. This is the crucial point that [12, 48] miss and is really only evident here because of the microphysical prospective. It does not matter if the pot is made of glass and transmits a large portion of light into the room, or made of terracotta and absorbs much of the light. In any case, EO cancellation must occur and the corresponding largeness of ρ results in this cancellation contributing a factor of G to the extinction cross section. The remaining factor of G then comes from the energy flow associated with the reflected (scattered) and/or absorbed light.

Notice that, within the photon particle-like view of scattering discussed in Sec. 1.7, only one factor of G of incident flux is scattered from the flower pot's illuminated side, and hence removed from the forward direction. It would then be inconsistent to count a contribution from the shadow; reflection from the pot's illuminated surface removes the light that would otherwise occupy the shadow. There is no need to fill the shadow with

"anti-particles" of light to cancel out the incident flux. Hence, this approach is not capable of accounting for the extra factor of G in the paradox. This example illustrates the inherent flaw of the "photon" concept and emphasizes the fundamental wave nature of Maxwell's theory. The incident wave permeates all space including the particle. Therefore, it must be a response of the particle that produces a wave that cancels the incident through destructive interference, resulting in the extinction paradox. This must occur regardless of the size of the wavelength. *It is the wave nature of light that accounts for an object's shadow*, not a mechanical-like blocking of the light.

Bibliography

[1] M. I. Mishchenko, L. D. Travis, and A. A. Lacis, *Multiple Scattering of Light by Particles*, Cambridge: Cambridge University Press, 2006.

[2] M. I. Mishchenko, L. D. Travis, and A. A. Lacis, *Scattering, Absorption, and Emission of Light by Small Particles*, Cambridge: Cambridge University Press, 2002, freely available on the `.pdf` format at http://www.giss.nasa.gov/ crmim/books.html.

[3] F. Hynne and R. Bullough, Phil. Trans. R. Soc. Lond. **312**, 251 (1984).

[4] F. K. Richtmeyer, E. H. Kennard, and J. N. Cooper, *Introduction to Modern Physics*, New York: McGraw-Hill, 1955.

[5] J. A. Stratton, *Electromagnetic Theory*, New York: McGraw-Hill, 1941.

[6] E. J. Rothwell and M. J. Cloud, *Electrodynamics*, New York: CRC Press, 2001.

[7] A. E. Siegman, *Lasers*, Sausalito: University Science Books, 1986.

[8] H. A. Lorentz, *The Theory of Electrons*, New York: Dover, 2003.

[9] C. Jeffries, J. Soc. Ind. Appl. Math **34**, 386 (1992).

[10] C. Jeffries, J. Soc. Ind. Appl. Math **36**, 638 (1992).

[11] R. H. Romer, Am. J. Phys. **50**, 1166 (1982).

[12] C. S. Lai, Am. J. Phys. **49**, 841 (1981).

[13] D. F. Nelson, Phys. Rev. Lett. **76**, 4713 (1996).

[14] G. J. J. Rikken and B. A. van Tiggelen, Phys. Rev. Lett. **78**, 847 (1997).

[15] J. D. Jackson, *Classical Electrodynamics*, New York: Wiley, 1999.

[16] W. Gough, Eur. J. Phys. **3**, 83 (1982).

[17] I. Campos and J. L. Jiménez, Eur. J. Phys. **13**, 117 (1992).

[18] P. C. Peters, Am. J. Phys. **50**, 1165 (1982).

[19] A. Chubykalo, A. Espinoza, and R. Tznochev, Eur. Phys. J. D **31**, 113 (2004).

[20] F. N. H. Robinson, J. Soc. Ind. Appl. Math **36**, 633 (1994).

[21] M. Nakajima, Electron. Commun. Jpn., Part 2: Electron. **82**, 56 (1999).

[22] D. J. Jackson, Am. J. Phys. **74**, 280 (2006).

[23] D. H. Kobe, Am. J. Phys. **50**, 1162 (1982).

[24] M. Born and E. Wolf, *Principles of optics*, Cambridge: Cambridge University Press, 1999.

[25] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, volume Vol. 2, Reading: Addison-Wesley Publishing Company, 1964.

[26] M. I. Mishchenko, Rev. Geophys. **46**, 2007RG00230 (2008).

[27] M. I. Mishchenko, Bull. Amer. Meteorol. Soc. , in press: doi:10.1175/2008BAMS2632.1 (2008).

[28] E. Merzbacher, *Quantum Mechanics*, volume Third Ed., New York: John Wiley & Sons, Inc., 1998.

[29] R. P. Feynman, *Quantum Electrodynamics*, New York: W. A. Benjamin, Inc., 1961.

[30] C.-T. Tai, *Dyadic Green Functions in Electromagnetic Theory*, volume Second Ed., New York: IEEE Press, 1994.

[31] P. C. Waterman, Proc. IEEE **53**, 805 (1965).

[32] L. Lewin, IEEE Trans. Microwave Theory Tech. **18**, 1041 (1970).

[33] J. van Bladel, *Singular Electromagnetic Fields and Sources*, Oxford: Oxford University Press, 1991.

[34] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, volume Vol. 1, Reading: Addison-Wesley Publishing Company, 1964.

[35] M. I. Mishchenko, M. J. Berg, C. M. Sorensen, and C. V. M. van der Mee, J. Quant. Spectros. Radiant. Transf. , In review (2009).

[36] W. T. Grandy, *Scattering of Waves by Large Spheres*, Cambridge: Cambridge University Press, 2000.

[37] C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, New York: Wiley, 1983.

[38] G. Videen and W. Sun, Appl. Opt. **42**, 6724 (2003).

[39] Q. Fu and W. Sun, Appl. Opt. **40**, 1354 (2001).

[40] M. I. Mishchenko, J. Quant. Spectros. Radiant. Transf. **109**, 2386 (2001).

[41] M. I. Mishchenko, Opt. Ex. **15**, 13188 (2007).

[42] M. I. Mishchenko, J. Quant. Spectros. Radiant. Transf. **101**, 404 (2006).

[43] M. I. Mishchenko, J. Quant. Spectros. Radiant. Transf. **100**, 268 (2006).

[44] A. Doicu, T. Wriedt, and Y. A. Eremin, *Light Scattering by Systems of Particles*, Berlin: Springer, 2006.

[45] M. Abramowitz and I. A. Stegun, *A Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, volume 9th Ed., Washington D.C.: National Bureau of Standards, 1970.

[46] G. Videen, R. G. Pinnick, D. Ngo, Q. Fu, and P. Chýlek, Appl. Opt. **37**, 1104 (1998).

[47] H. C. van de Hulst, *Light Scattering by Small Particles*, New York: Dover, 1981.

[48] L. Brillouin, J. Appl. Phys. **20**, 1110 (1949).

[49] O. Heaviside, *Electromagnetic Theory*, volume Vol. 1, London: The Electrician Printing & Publishing Co., 1893.

[50] L. Tsang, J. A. Kong, K.-H. Ding, and C. O. Ao, *Scattering of Electromagnetic Waves: Numerical Simulations*, New York: John Wiley & Sons, Inc., 2001.

[51] M. A. Yurkin and A. G. Hoekstra, J. Quant. Spectros. Radiant. Transf. **106**, 558 (2007).

[52] E. M. Purcell and C. R. Pennypacker, Astrophys. J. **186**, 705 (1973).

[53] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, volume Vol. 3, Reading: Addison-Wesley Publishing Company, 1964.

[54] R. K. Pathria, *Statistical Mechanics*, volume 2nd Ed., Oxford: Butterworth-Heinemann, 2003.

[55] N. G. Khlebstov, Opt. Spectros. **90**, 408 (2000).

[56] D. W. Mackowski, J. Opt. Soc. Am. A **19**, 881 (2002).

[57] A. Morawiec, *Orientations and Rotations*, Berlin: Springer, 2004.

[58] D. J. Griffiths, *Introduction to Electrodynamics*, volume 3rd Ed., Upper Saddle River: Prentice-Hall, 1999.

[59] E. Hecht, *Optics*, volume 4th Ed., New York: Addison-Wesley, 2004.

[60] C. E. Baum and N. H. Kritikos, *Electromagnetic Symmetry*, Washington D.C.: Taylor & Francis, 1995.

[61] P. C. Waterman, Phys. Rev. D **3**, 825 (1971).

[62] J. W. Hovenier, J. Atmpos. Sci. **26**, 488 (1969).

[63] J. W. Hovenier, J. Astron. Astrophys. **7**, 86 (1970).

[64] S. V. Nghiem, S. H. Yueh, R. Kwok, and F. K. Li, Radio Sci. **27**, 693 (1992).

[65] F. M. Schultz, K. Stamnes, and J. J. Stamnes, J. Opt. Soc. Am. A **16**, 853 (1999).

[66] J. W. Hovenier, J. Quant. Spectros. Radiant. Transf. **60**, 483 (1998).

[67] S. H. Yueh, R. Kwok, and S. V. Nghiem, Radio Sci. **29**, 1409 (1994).

[68] K. F.Ren, G. Gréhan, and G. Gouesbet, J. Opt. Soc. Am. A **11**, 1812 (1994).

[69] C. H. Hu, G. W. Kattawar, M. E. Parkin, and P. Herb, Appl. Opt. **26**, 4159 (1987).

[70] J. Tyynelä, E. Zubko, G. Videen, and K. Muinonen, J. Quant. Spectros. Radiant. Transf. **106**, 520 (2007).

[71] B. Chu, *Laser Light Scattering*, volume 2nd Ed., New York: Dover, 2007.

[72] J. F. Nye, Proc. R. Soc. London, Ser. A **389**, 279 (1983).

[73] J. F. Nye, Philos. Trans. R. Soc. London, Ser. A **355**, 2065 (1997).

[74] C. M. Sorensen, Aerosol. Sci. Tech. **35**, 648 (2001).

[75] C. M. Sorensen and D. J. Fishbach, Opt. Commun. **173**, 145 (2000).

[76] M. Kerker, *The Scattering of Light and Other Electromagnetic Radiation*, New York: Academic Press, 1969.

[77] C. M. Sorensen and D. Shi, Opt. Commun. **178**, 31 (2000).

[78] J. C. Jonsson, G. B. Smith, and G. A. Niklasson, Opt. Commun. **240**, 9 (2004).

[79] M. J. Berg, C. M. Sorensen, and A. Chakrabarti, Appl. Opt. **44**, 7487 (2004).

[80] A. Guinier and G. Fournet, *Small-Angle Scattering of X-Rays*, New York: John Wiley & Sons, Inc., 1955.

[81] T. C. Bond and R. W. Bergstrom, Aerosol. Sci. Tech. **40**, 27 (2006).

[82] L. Liu and M. I. Mishchenko, J. Geophys. Res. **110**, D11211 (2005).

[83] R. Dhaubhadel, *An Experimental Study of Dense Aerosol Aggregations*, PhD thesis, Kansas State University, 2008.

[84] F. G.Pierce, *Aggregation in Colloids and Aerosols*, PhD thesis, Kansas State University, 2007.

[85] J. Feder, *Fractals*, New York: Plenum, 1988.

[86] G. Wang and C. M. Sorensen, Appl. Opt. **41**, 4645 (2002).

[87] N. Lu and C. M. Sorensen, Phys. Rev. E **50**, 3109 (1994).

[88] R. Dhaubhadel, F. Pierce, A. Chakrabarti, and C. M. Sorensen, Phys. Rev. E **73**, 011404 (2006).

[89] R. E. Collin, *Field Theory of Guided Waves*, volume 2nd Ed., New York: IEEE Press, 1990.

[90] L. Tsang, J. A. Kong, K.-H. Ding, and C. O. Ao, *Scattering of Electromagnetic Waves: Theories and Applications*, New York: John Wiley & Sons, Inc., 2001.

[91] G. S. Smith, *An Introduction to Electromagnetic Radiation*, Cambridge: Gambridge University Press, 1997.

[92] R. G. Newton, Phys. Rev. E **50**, 3109 (1994).

[93] R. G. Newton, *Scattering Theory of Waves and Particles*, New York: Dover, 2002.

[94] P. S. Carney, J. C. Schotland, and E. Wolf, Phys. Rev. E **70**, 036611 (2004).

[95] A. A. Kokhanovsky, *Light Scattering Reviews*, Chichester: Springer, Praxis Publishing, 2006.

[96] P. Chýlek, J. Opt. Soc. Am. **67**, 561 (1977).

[97] J. A. Lock, J. T. Hodges, and G. Gouesbet, J. Opt. Soc. Am. A **12**, 2708 (1995).

[98] H. C. van de Hulst, Physica (Amsterdam) **15**, 740 (1949).

[99] C.-T. Tai, J. Electromagn. Waves Appl. **16**, 597 (2002).

[100] J. A. Grzesik, Prog. Electromagn. Res. **40**, 255 (2003).

[101] M. I. Mishchenko, J. W. Hovenier, and D. W. Mackowski, Prog. Electromagn. Res. **40**, 255 (2003).

[102] P. A. Martin, *Multiple Scattering: Interaction of Time-Harmonic Waves with N Particles*, Cambridge: Cambridge University Press, 2006.

[103] S. Asano, Appl. Opt. **18**, 712 (1979).

[104] H. M. Lai, W. Y. Wong, and W. H. Wong, J. Opt. Soc. Am. A **21**, 2324 (2004).

[105] A. Lock and L. Yang, J. Opt. Soc. Am. A **8**, 1132 (1991).

[106] T. H. Barr, *Vector Calculus*, Upper Saddle River: Prentice Hall, 1997.

[107] D. W. J. Cruickshank, *P. P. Ewald and his Dynamical Theory of X-Ray Scattering*, Oxford: Oxford University Press, 1992.

[108] R. W. P. King and T. T. Wu, *The Scattering and Diffraction of Waves*, Cambridge, Massachusetts: Harvard University Press, 1959.

[109] R. Harrington, IRE Trans. Antennas Prop. , 150 (1959).

[110] J. J. Sakurai, *Modern Quantum Mechanics*, New York: Addison Wesley Longman, 1994.

[111] D. A. H. Jacobs, IMA J. Numerical Analysis **6**, 447 (1986).

[112] G. B. Arfken and H. J. Webber, *Mathematical Methods for Physicists*, San Diego: Harcourt, 2001.

Appendix A

Various Vector and Dyadic Relations

This appendix provides a brief list of relations that are needed for the calculations in this work.

$$\left[\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\overleftrightarrow{\mathbf{A}}(\mathbf{r},\mathbf{r}')\right]\cdot\mathbf{b}(\mathbf{r}')=\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\left[\overleftrightarrow{\mathbf{A}}(\mathbf{r},\mathbf{r}')\cdot\mathbf{b}(\mathbf{r}')\right] \tag{A.1}$$

$$\boldsymbol{\nabla}\times\left[\overleftrightarrow{\mathbf{I}}f(\mathbf{r},\mathbf{r}')\right]=[\boldsymbol{\nabla}f(\mathbf{r},\mathbf{r}')]\times\overleftrightarrow{\mathbf{I}} \tag{A.2}$$

$$\boldsymbol{\nabla}\left[\frac{\exp(ikr)}{r}\right]=\left(ik-\frac{1}{r}\right)\frac{\exp(ikr)}{r}\,\hat{\mathbf{r}}, \tag{A.3}$$

$$\left(\hat{\mathbf{r}}\times\overleftrightarrow{\mathbf{I}}\right)\cdot\mathbf{a}=\hat{\mathbf{r}}\times\mathbf{a} \tag{A.4}$$

$$\hat{\mathbf{r}}\times(\hat{\mathbf{r}}\otimes\hat{\mathbf{r}})=0 \tag{A.5}$$

$$(\hat{\mathbf{r}}\otimes\hat{\mathbf{r}})\cdot\mathbf{a}=\hat{\mathbf{r}}\cdot\mathbf{a} \tag{A.6}$$

Appendix B

Various Properties of Vector Spherical Wavefunctions

The following properties of the VSWFs and the vector spherical harmonics (VSHs) can be found in [2, App. C] but are repeated here for completeness and reference.

The curl relationships between the VSWF and the regular VSWF are

$$\mathbf{M}_{mn}(k\mathbf{r}) = \frac{1}{k}\nabla \times \mathbf{N}_{mn}(k\mathbf{r}), \tag{B.1}$$

$$\mathbf{N}_{mn}(k\mathbf{r}) = \frac{1}{k}\nabla \times \mathbf{M}_{mn}(k\mathbf{r}), \tag{B.2}$$

$$\mathrm{Rg}\mathbf{M}_{mn}(k\mathbf{r}) = \frac{1}{k}\nabla \times \mathrm{Rg}\mathbf{N}_{mn}(k\mathbf{r}), \tag{B.3}$$

$$\mathrm{Rg}\mathbf{N}_{mn}(k\mathbf{r}) = \frac{1}{k}\nabla \times \mathrm{Rg}\mathbf{M}_{mn}(k\mathbf{r}). \tag{B.4}$$

The VSWFs can be expressed in terms of products of the spherical Bessel and Hankel functions of the first kind j_n and $h_n^{(1)}$, respectively, and the VSHs, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{P}$ as

$$\mathbf{M}_{mn}(k\mathbf{r}) = \gamma_{mn} h_n^{(1)}(kr)\mathbf{C}_{mn}(\hat{\mathbf{r}}), \tag{B.5}$$

$$\mathrm{Rg}\mathbf{M}_{mn}(k\mathbf{r}) = \gamma_{mn} j_n(kr)\mathbf{C}_{mn}(\hat{\mathbf{r}}), \tag{B.6}$$

$$\mathbf{N}_{mn}(k\mathbf{r}) = \gamma_{mn}\Big\{\frac{n(n+1)}{kr} h_n^{(1)}(kr)\mathbf{P}_{mn}(\hat{\mathbf{r}}) + \frac{1}{kr}\big[kr\, h_n^{(1)}(kr)\big]'\mathbf{B}_{mn}(\hat{\mathbf{r}})\Big\}, \tag{B.7}$$

$$\mathrm{Rg}\mathbf{N}_{mn}(k\mathbf{r}) = \gamma_{mn}\Big\{\frac{n(n+1)}{kr} j_n(kr)\mathbf{P}_{mn}(\hat{\mathbf{r}}) + \frac{1}{kr}\big[kr\, j_n(kr)\big]'\mathbf{B}_{mn}(\hat{\mathbf{r}})\Big\}, \tag{B.8}$$

where

$$\gamma_{mn} = \sqrt{\frac{(2n+1)(n-m)!}{4\pi n(n+1)(n+m)!}}. \tag{B.9}$$

The VSHs are interrelated as

$$\mathbf{B}_{mn}(\hat{\mathbf{r}}) = \hat{\mathbf{r}} \times \mathbf{C}_{mn}(\hat{\mathbf{r}}), \tag{B.10}$$

$$\mathbf{C}_{mn}(\hat{\mathbf{r}}) = -\hat{\mathbf{r}} \times \mathbf{B}_{mn}(\hat{\mathbf{r}}), \tag{B.11}$$

with the additional property

$$\hat{\mathbf{r}} \times \mathbf{P}_{mn}(\hat{\mathbf{r}}) = 0. \tag{B.12}$$

The orthogonality relationships for the VSHs are

$$\int_{4\pi} \mathbf{B}_{mn}(\hat{\mathbf{r}}) \cdot \mathbf{C}^*_{m'n'}(\hat{\mathbf{r}})\, \mathrm{d}\Omega = \int_{4\pi} \mathbf{B}^*_{mn}(\hat{\mathbf{r}}) \cdot \mathbf{C}_{m'n'}(\hat{\mathbf{r}})\, \mathrm{d}\Omega = 0, \tag{B.13}$$

$$\int_{4\pi} \mathbf{B}_{mn}(\hat{\mathbf{r}}) \cdot \mathbf{B}^*_{m'n'}(\hat{\mathbf{r}})\, \mathrm{d}\Omega = \int_{4\pi} \mathbf{C}^*_{mn}(\hat{\mathbf{r}}) \cdot \mathbf{C}_{m'n'}(\hat{\mathbf{r}})\, \mathrm{d}\Omega = \frac{\delta_{mm'}\delta_{nn'}}{(\gamma_{mn})^2}. \tag{B.14}$$

With the orthogonality properties above, one can derive the expansion coefficients a_{mn} and b_{mn} for the incident plane wave of Eqs. (2.19), (2.22), (3.1), and (3.2) as

$$a_{mn} = 4\pi(-1)^m i^n d_n \mathbf{E}^{inc}_o \cdot [\mathbf{C}_{mn}]^*, \tag{B.15}$$

$$b_{mn} = 4\pi(-1)^m i^{n-1} d_n \mathbf{E}^{inc}_o \cdot [\mathbf{B}_{mn}]^*, \tag{B.16}$$

where

$$d_n = \sqrt{\frac{2n+1}{4\pi n(n+1)}}. \tag{B.17}$$

Appendix C

Complex Bi-Conjugate Gradient Solution Method

Multiple methods are known for the solution of matrix equations like Eq. (5.15). An efficient iterative method is the Complex Bi-Conjugate Gradient Method (CBCGM). This method is well suited for solving Eq. (5.15) because it solves the system without needing the explicit values of elements of $\bar{\bar{\mathbf{A}}}$ or $\bar{\mathbf{p}}$; only terms of the matrix vector products $\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}$ are needed [111].

To describe the CBCGM method, let

$$\bar{\mathbf{b}} = \bar{\mathbf{E}}^{inc}, \quad \bar{\mathbf{x}} = \bar{\mathbf{p}}, \tag{C.1}$$

and define an initial guess column vector $\bar{\mathbf{x}}_o$ as

$$\bar{\mathbf{x}}_o = \bar{\mathbf{E}}^{inc}. \tag{C.2}$$

The choice for the guess in Eq. (C.2) is known as Jacobi preconditioning and is not arbitrary and is intended to lead to faster convergence to the solution, see [51]. Next the initial residual column vector $\bar{\mathbf{r}}_o$ and initial bi-residual column vector $\tilde{\mathbf{r}}_o$ are calculated

$$\bar{\mathbf{r}}_o = \bar{\mathbf{b}} - \bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{x}}_o, \quad \tilde{\mathbf{r}}_o = \bar{\mathbf{r}}_o^*, \tag{C.3}$$

then the initial search direction column vector $\bar{\mathbf{p}}_o$ and initial bi-direction column vector $\tilde{\mathbf{p}}_o$

$$\bar{\mathbf{p}}_o = \bar{\mathbf{r}}_o, \quad \tilde{\mathbf{p}}_o = \bar{\mathbf{p}}_o^*. \tag{C.4}$$

The dimension of the coefficient matrix $\bar{\bar{\mathbf{A}}}$ is $3N$ and the CBCGM is designed to converge to the solution of Eq. (5.15) in at most $3N-1$ iterations. To keep track of each iteration, the index i is introduced. Then, for $\mathsf{i}=0,\ldots 3N-1$ one calculates the following:

- The step length parameter α_{i}

$$\alpha_{\mathsf{i}} = \frac{\bar{\mathbf{r}}_{\mathsf{i}}^{*}\cdot\bar{\mathbf{r}}_{\mathsf{i}}}{\bar{\mathbf{p}}_{\mathsf{i}}^{*}\cdot\left(\bar{\bar{\mathbf{A}}}\cdot\bar{\mathbf{p}}_{\mathsf{i}}\right)}. \tag{C.5}$$

- The new solution is then given by

$$\bar{\mathbf{x}}_{\mathsf{i}+1} = \bar{\mathbf{x}} + \alpha_{\mathsf{i}}\bar{\mathbf{p}}_{\mathsf{i}}, \tag{C.6}$$

- The next residual and bi-residual are

$$\bar{\mathbf{r}}_{\mathsf{i}+1} = \bar{\mathbf{r}}_{\mathsf{i}} - \alpha_{\mathsf{i}}\bar{\bar{\mathbf{A}}}\cdot\bar{\mathbf{p}}_{\mathsf{i}}, \quad \bar{\mathbf{r}}_{\mathsf{i}+1} = \bar{\mathbf{r}}_{\mathsf{i}} - \alpha_{\mathsf{i}}^{*}\bar{\bar{\mathbf{A}}}^{\dagger}\cdot\bar{\mathbf{p}}_{\mathsf{i}}^{*}, \tag{C.7}$$

- Calculate the bi-conjugacy coefficient

$$\beta_{\mathsf{i}} = -\frac{\left(\bar{\bar{\mathbf{A}}}^{\dagger}\cdot\bar{\mathbf{p}}_{\mathsf{i}}^{*}\right)^{\dagger}\cdot\bar{\mathbf{r}}_{\mathsf{i}+1}}{\bar{\mathbf{p}}_{\mathsf{i}}^{*}\cdot\left(\bar{\bar{\mathbf{A}}}\cdot\bar{\mathbf{p}}_{\mathsf{i}}\right)} \tag{C.8}$$

- Then the next direction and bi-direction vectors are

$$\bar{\mathbf{p}}_{\mathsf{i}+1} = \bar{\mathbf{r}}_{\mathsf{i}+1} + \beta_{\mathsf{i}}\bar{\mathbf{p}}_{\mathsf{i}}, \quad \bar{\mathbf{p}}_{\mathsf{i}+1} = \bar{\mathbf{r}}_{\mathsf{i}+1} + \beta_{\mathsf{i}}^{*}\bar{\mathbf{p}}_{\mathsf{i}}^{*}. \tag{C.9}$$

After each iteration of the CBCGM, the column vector $\bar{\mathbf{x}}$ approaches the solution of the matrix system from the initial guess $\bar{\mathbf{x}}_o$. In most cases the algorithm will converge in far less than $3N-1$ iterations, which is fortunate since a typical value for N may be more than ten million. The convergence of the algorithm can be monitored by comparing the sum of all the elements of the residual column vector $\bar{\mathbf{r}}$ to a tolerance δ. If the sum is less than the tolerance, iterations can be stopped and the resulting solution for the dipole moments will be correct to within a maximum error of δ.

Another issue with the CBCGM concerns how one accounts for an arbitrary particle shape. The initial guess column vector $\bar{\mathbf{x}}_o$ and the solution column vector $\bar{\mathbf{x}}$ contain the dipole moments of the *entire* cubic lattice array. This effectively makes the particle geometry a cube with the same dimensions as the lattice $\mathfrak{n}$. Consequently, if the particle shape is not cubic, the dipole moments located outside the particle volume are set to zero. This is done at each step described above where the column vectors $\bar{\mathbf{x}}_o$, $\bar{\mathbf{r}}_o$, $\bar{\mathbf{x}}$ and $\bar{\mathbf{r}}$ are involved.

Appendix D

Matrix Multiplication using Discrete Fourier Transforms

The CBCGM described in App. C is an efficient and fast route to find the solution to Eq. (5.15). However, it requires terms of the matrix-vector products $\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}_{\mathfrak{i}}$ and $\bar{\bar{\mathbf{A}}}^{\dagger} \cdot \bar{\mathbf{p}}_{\mathfrak{i}}^{*}$. This is a problem, because, evaluating a matrix-vector product directly would be a $(3N)^2$ operation and hence would be computationally impractical for even moderately large N. So, in order to make practical use of the CBCGM, a fast and low memory-consuming way to find matrix-vector products is needed.

The properties of the coefficient matrix discussed in Sec. 5.2 show that the matrix equation (5.15) is equivalent to the discrete convolution of Eq. (5.19). A discrete convolution can be evaluated by calculating using the Fast Fourier Transform (FFT), which scales in computational time as $\sim N \log N$ rather than $(3N)^3$ as would be the case for direct evaluation of the convolution. This section will briefly describe the Discrete Fourier Transform (DFT) and how it is used to evaluate FFT-based Toeplitz matrix multiplication.

The Fourier transform of a function $f(x)$ can be expressed as

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} \mathrm{d}x, \tag{D.1}$$

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) e^{ikx} \mathrm{d}k, \tag{D.2}$$

where the variable k is the spatial frequency associated with the variable x. To extend Eqs. (D.1) and (D.2) to the discrete case, consider a total of $\mathfrak{n}$ samples of the continuous function

$f(x)$ at constant intervals d,

$$f(x) \to f(x_n), \quad f_n \equiv f(x_n), \quad \text{where} \quad x_n = d, 2d, 3d \ldots \mathfrak{n}d. \tag{D.3}$$

For convenience let $d = 1$, then the set of sampled values forms the sequence,

$$f_n \in \{f(1), f(2), f(3), \ldots f(\mathfrak{n})\}.$$

Following [50, Sec. 4.1], define a complex number

$$W_{\mathfrak{n}} = \exp\left(\frac{i2\pi}{\mathfrak{n}}\right), \tag{D.4}$$

then the DFT of the sequence f_n is

$$F(k) = \sum_{n=1}^{\mathfrak{n}} f(n) W_{\mathfrak{n}}^{(k-1)(n-1)}, \tag{D.5}$$

where $\{k, n\} \in \{1, 2, 3, \ldots \mathfrak{n}\}$ and the inverse DFT is

$$f(n) = \frac{1}{\mathfrak{n}} \sum_{k=1}^{\mathfrak{n}} F(k) W_{\mathfrak{n}}^{-(k-1)(n-1)}. \tag{D.6}$$

The goal is to evaluate an equation like

$$y(n) = \sum_{m=1}^{N} a(n-m) f(n), \quad \{n, m\} \in \{1, 2, 3, \ldots, \mathfrak{n}\}, \tag{D.7}$$

using the convolution theorem and DFTs. This is *not* as straight forward as simply evaluating the DFTs of $a(x)$ and $f(x)$, multiplying the resultant transforms and then evaluating the inverse transform [50, p. 37]. To see why, consider the number of points at which the function $a(x)$ is sampled. If $\{n, m\} \in \{1, 2, 3, \ldots \mathfrak{n}\}$ then $a(n-m)$ is sampled at $a(-\mathfrak{n}+1), \ldots a(0), \ldots, a(\mathfrak{n}-1)$, which is a total of $2\mathfrak{n}-1$ sample points. This is a complication because the sequences $a(n)$ and $f(n)$ are unequal in length. This problem is resolved with zero padding. Let,

$$f_1(n) = \begin{cases} f(n), & \text{for } n = 1, 2, 3, \ldots \mathfrak{n} \\ 0, & \text{for } n = \mathfrak{n}+1, \mathfrak{n}+2, \mathfrak{n}+3, \ldots 2\mathfrak{n}. \end{cases} \tag{D.8}$$

If $f_1(n)$ is used in place of $f(n)$ then nothing is changed in Eq. (D.7) provided that the limit on the sum is changed form $\mathfrak{n}$ to $2\mathfrak{n}$ so that the entire length of the sequence $a(n-m)$ is included.

Now the convolution theorem can be used. Let $A(k)$ and $F(k)$ be the DFTs of $a(n)$ and $f(n)$, respectively, and let $y(n)$ be given by Eq. (D.7) with transform $Y(k)$. The convolution theorem states that

$$Y(k) = A(k)F(k). \tag{D.9}$$

Using Eq. (D.5) and (D.6), $y(n)$ is given by

$$\begin{aligned}
y(n) &= \frac{1}{2\mathfrak{n}} \sum_{k=1}^{2\mathfrak{n}} Y(k) W_{\mathfrak{n}}^{-(k-1)(n-1)} \\
&= \frac{1}{2\mathfrak{n}} \sum_{k=1}^{2\mathfrak{n}} A(k)F(k) W_{\mathfrak{n}}^{-(k-1)(n-1)} \\
&= \frac{1}{2\mathfrak{n}} \sum_{k=1}^{2\mathfrak{n}} \left[\sum_{m=1}^{2\mathfrak{n}} a(m) W_{\mathfrak{n}}^{(k-1)(m-1)} \right] \left[\sum_{r=1}^{2\mathfrak{n}} f(r) W_{\mathfrak{n}}^{(k-1)(r-1)} \right] W_{\mathfrak{n}}^{-(k-1)(n-1)} \\
&= \frac{1}{2\mathfrak{n}} \sum_{k,m,r=1}^{2\mathfrak{n}} a(m)f(r) W_{\mathfrak{n}}^{(k-1)(m-1)+(k-1)(r-1)-(k-1)(n-1)} \\
&= \frac{1}{2\mathfrak{n}} \sum_{k,m,r=1}^{2\mathfrak{n}} a(m)f(r) W_{\mathfrak{n}}^{(k-1)(m+r-n-1)} \\
&= \sum_{n,r=1}^{2\mathfrak{n}} a(n-r+1)f(r),
\end{aligned} \tag{D.10}$$

where in the last step of Eq. (D.10) the following orthogonality condition is used [112, p. 898]

$$\frac{1}{2\mathfrak{n}} \sum_{k=1}^{2\mathfrak{n}} W_{2\mathfrak{n}}^{-(k-1)m} \left[W_{2\mathfrak{n}}^{-(k-1)(n-r+1)} \right]^* = \delta_{m,n-r+1}. \tag{D.11}$$

Now the problem is that Eq. (D.7), which is needed to evaluate a matrix-vector product, is not in the same form as the discrete convolution appearing in Eq. (D.10). In the steps leading to (D.10), the sequence $a(n)$ gets shifted by one interval. If the sequence $a(n)$ is regarded as periodic, then this shifting moves part of one period of the sequence into the

next. This can be resolved by by defining a new sequence from $g(n)$ [50, Sec. 4.1]

$$a_1(n) = \begin{cases} a(n-1) & \text{for, } 1 \leqslant n \leqslant \mathfrak{n} \\ 0 & \text{for, } n = \mathfrak{n}+1 \\ a(n-2N-1) & \text{for, } \mathfrak{n}+2 \leqslant n \leqslant 2\mathfrak{n}, \end{cases} \tag{D.12}$$

then the convolution equation (D.7) becomes

$$y(n) = \sum_{m=1}^{2\mathfrak{n}} a_1(n-m+1) f_1(n), \tag{D.13}$$

which is now in the same form as Eq. (D.10). The following will discuss the application of these results to the task of evaluating a Toeplitz matrix-vector product.

To demonstrate the method for evaluating a matrix-vector product using DFTs, consider a square Toeplitz matrix $\bar{\bar{\mathbf{Q}}}$ of dimension $N \times N$ with known elements q_{nm} given by a kernel $q(n,m)$,

$$\bar{\bar{\mathbf{Q}}} = \begin{pmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1N} \\ q_{21} & q_{22} & q_{23} & \cdots & q_{2N} \\ q_{31} & q_{32} & q_{33} & \cdots & q_{3N} \\ \vdots & \vdots & \vdots & & \vdots \\ q_{N1} & q_{N2} & q_{N3} & \cdots & q_{NN} \end{pmatrix}.$$

Because $\bar{\bar{\mathbf{Q}}}$ is Toeplitz, its elements depend only on the difference of their indices, so

$$\bar{\bar{\mathbf{Q}}} = \begin{pmatrix} q_0 & q_{-1} & q_{-2} & \cdots & q_{1-N} \\ q_1 & q_0 & q_{-1} & \cdots & q_{2-N} \\ q_2 & q_1 & q_0 & \cdots & q_{3-N} \\ \vdots & \vdots & \vdots & & \vdots \\ q_{N-1} & q_{N-2} & q_{N-3} & \cdots & q_0 \end{pmatrix}.$$

Let an arbitrary column vector $\bar{\mathbf{x}}$ be given by

$$\bar{\mathbf{x}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{pmatrix},$$

then, the matrix vector product $\bar{\bar{\mathbf{Q}}} \cdot \bar{\mathbf{x}}$ will be another column vector $\bar{\mathbf{y}}$,

$$\bar{\mathbf{y}} = \bar{\bar{\mathbf{Q}}} \cdot \bar{\mathbf{x}} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_N \end{pmatrix}.$$

The steps involved to evaluate $\bar{\mathbf{y}}$ are as follows:

- From the elements of $\bar{\bar{\mathbf{Q}}}$ and $\bar{\mathbf{x}}$, form the sequences

$$q^{(1)}(n) = \begin{cases} q(n-1) & \text{for, } 1 \leqslant n \leqslant N \\ 0 & \text{for, } n = N+1 \\ q(n-2N-1) & \text{for, } N+2 \leqslant n \leqslant 2N, \end{cases}$$
$$x^{(1)}(n) = \begin{cases} x(n) & \text{for, } n = 1, 2, 3, \ldots N \\ 0 & \text{for, } n = N+1, N+2, N+3, \ldots 2N. \end{cases} \tag{D.14}$$

- Evaluate the discrete Fourier transforms of the sequences $q^{(1)}$ and $x^{(1)}$,

$$Q^{(1)}(k) = \sum_{n=1}^{2N} q_n^{(1)} W_{2N}^{(k-1)(n-1)}, \tag{D.15}$$
$$X^{(1)}(k) = \sum_{n=1}^{2N} x_n^{(1)} W_{2N}^{(k-1)(n-1)}. \tag{D.16}$$

- Multiply the transforms to form a new sequence

$$X^{(2)}(k) = Q^{(1)}(k) X^{(1)}(k) \tag{D.17}$$
$$= \left\{ Q_1^{(1)} X_1^{(1)}, Q_2^{(1)} X_2^{(1)}, \ldots Q_{2N}^{(1)} X_{2N}^{(1)} \right\}. \tag{D.18}$$

- Take the inverse transform of $X^{(2)}(k)$

$$x^{(2)}(n) = \frac{1}{2N} \sum_{k=1}^{2N} X_k^{(2)} W_{2N}^{-(k-1)(n-1)}. \tag{D.19}$$

- Form a sequence from the from the first N terms of $x^{(2)}(n)$

$$x^{(3)}(n) = \left\{ x_1^{(2)}, x_2^{(2)}, \ldots x_N^{(2)} . \right\} \tag{D.20}$$

- Then, the matrix vector product $\bar{\mathbf{y}}$, is equal to the column vector formed by $x^{(3)}(n)$

$$\bar{\mathbf{y}} = \begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \\ \vdots \\ x_N^{(3)} \end{pmatrix}. \tag{D.21}$$

Now recall that Eq. (5.19) involves a coefficient matrix $\bar{\bar{\mathbf{A}}}$ formed by a dyadic kernel with three indices $\{n, m, l\}$, as opposed to the single index n above

$$\mathbf{E}^{inc}(\mathbf{r}_{nml}) = \sum_{n',m',l'=1}^{N} \bar{\bar{\mathbf{A}}}_{n-n',m-m',l-l'} \cdot \mathbf{p}_{n'm'l'}.$$

Writing this out explicitly,

$$\begin{pmatrix} E_x^{inc}(n,m,l) \\ E_y^{inc}(n,m,l) \\ E_z^{inc}(n,m,l) \end{pmatrix} = \sum_{\substack{n'=1 \\ m'=1 \\ l'=1}}^{N} \left\{ \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \cdot \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} \right\}, \tag{D.22}$$

where

$$\begin{aligned} a_{ij} &= a_{ij}(n-n', m-m', l-l') && i,j \in \{1,2,3\}, \\ p_\gamma &= p_\gamma(n', m', l') && \gamma \in \{x,y,z\}. \end{aligned} \tag{D.23}$$

This reveal that Eq. (5.19) separates into three equations, one for each Cartesian component,

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_x = \sum_{n',m',l'=1}^{N} \left(a_{11}p_x + a_{12}p_y + a_{13}p_z\right) \hat{\mathbf{x}}, \tag{D.24}$$

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_y = \sum_{n',m',l'=1}^{N} \left(a_{21}p_x + a_{22}p_y + a_{23}p_z\right) \hat{\mathbf{y}}, \tag{D.25}$$

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_z = \sum_{n',m',l'=1}^{N} \left(a_{31}p_x + a_{32}p_y + a_{33}p_z\right) \hat{\mathbf{z}}. \tag{D.26}$$

Let $\circledast$ denote the convolution operation, then Eqs. (D.24)-(D.26) become,

$$\begin{aligned} \left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_x &= a_{11} \circledast p_x + a_{12} \circledast p_y + a_{13} \circledast p_z, \\ \left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_y &= a_{21} \circledast p_x + a_{22} \circledast p_y + a_{23} \circledast p_z, \\ \left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_x &= a_{31} \circledast a_x + g_{32} \circledast a_y + a_{33} \circledast p_z. \end{aligned}$$

Let the DFT of a sequence $f(n)$ be given by $\mathrm{DFT}\{f(n)\}$ and the inverse transform by $\mathrm{DFT}^{-1}\{f(n)\}$, then

$$A_{ij} = \mathrm{DFT}\{a_{ij}\}, \tag{D.27}$$

$$P_\gamma = \mathrm{DFT}\{p_\gamma\}, \tag{D.28}$$

and from the convolution theorem Eqs. (D.24)-(D.26) are,

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_x = \mathrm{DFT}^{-1}\left\{A_{11}P_x + A_{12}P_y + A_{13}P_z\right\} \tag{D.29}$$

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_y = \mathrm{DFT}^{-1}\left\{A_{21}P_x + A_{22}P_y + A_{23}P_z\right\} \tag{D.30}$$

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_z = \mathrm{DFT}^{-1}\left\{A_{31}P_x + A_{32}P_y + A_{33}P_z\right\}. \tag{D.31}$$

Because the convolution operation is linear,

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_x = \mathrm{DFT}^{-1}\left\{A_{11}P_x\right\} + \mathrm{DFT}^{-1}\left\{A_{12}P_y\right\} + \mathrm{DFT}^{-1}\left\{A_{13}P_z\right\}, \tag{D.32}$$

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_y = \mathrm{DFT}^{-1}\left\{A_{21}P_x\right\} + \mathrm{DFT}^{-1}\left\{A_{22}P_y\right\} + \mathrm{DFT}^{-1}\left\{A_{23}P_z\right\}, \tag{D.33}$$

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_z = \mathrm{DFT}^{-1}\left\{A_{31}P_x\right\} + \mathrm{DFT}^{-1}\left\{A_{32}P_y\right\} + \mathrm{DFT}^{-1}\left\{A_{33}P_z\right\}. \tag{D.34}$$

Now all that is needed is a method to evaluate the discrete forward and inverse transforms of sequences that depend on three indices. This is done by a direct extension of the one index case. First, each sequence is padded with zeros;

$$s_x(n,m,l) = \begin{cases} p_x(n,m,l) & \begin{array}{c} 1 \leqslant n \leqslant N, \quad 1 \leqslant m \leqslant N \\ 1 \leqslant l \leqslant N \end{array} \\ 0 & \begin{array}{c} N+1 \leqslant n \leqslant 2N, \quad N+1 \leqslant m \leqslant 2N \\ N+1 \leqslant l \leqslant 2N, \end{array} \end{cases} \tag{D.35}$$

$$s_y(n,m,l) = \begin{cases} p_y(n,m,l) & \begin{array}{c} 1 \leqslant n \leqslant N, \quad 1 \leqslant m \leqslant N \\ 1 \leqslant l \leqslant N \end{array} \\ 0 & \begin{array}{c} N+1 \leqslant n \leqslant 2N, \quad N+1 \leqslant m \leqslant 2N \\ N+1 \leqslant l \leqslant 2N, \end{array} \end{cases} \tag{D.36}$$

$$s_z(n,m,l) = \begin{cases} p_z(n,m,l) & \begin{array}{c} 1 \leqslant n \leqslant N, \quad 1 \leqslant m \leqslant N \\ 1 \leqslant l \leqslant N \end{array} \\ 0 & \begin{array}{c} N+1 \leqslant n \leqslant 2N, \quad N+1 \leqslant m \leqslant 2N \\ N+1 \leqslant l \leqslant 2N, \end{array} \end{cases} \tag{D.37}$$

and let $a_{ij}(n,m,l) = 0$ for $i,j \in \{1,2,3\}$ unless [50, p. 46],

Then, the discrete Fourier transforms of the new sequences are:

$$S_\gamma(k_n,k_m,k_l) = \sum_{n,m,l=1}^{2N} s_\gamma(n,m,l) W_{2N}^{(k_n-1)(n-1)} W_{2N}^{(k_m-1)(m-1)} W_{2N}^{(k_l-1)(l-1)}, \tag{D.38}$$

n	m	l	$g_{ij}(n,m,l)$
$[1,N]$	$[1,N]$	$[1,N]$	$a_{ij}(n-1,m-1,l-1)$
$[N+2,2N]$	$[1,N]$	$[1,N]$	$a_{ij}(n-2N-1,m-1,l-1)$
$[1,N]$	$[N+2,2N]$	$[1,N]$	$a_{ij}(n-1,m-2N-1,l-1)$
$[N+2,2N]$	$[N+2,2N]$	$[1,N]$	$a_{ij}(n-2N-1,m-2N-1,l-1)$
$[1,N]$	$[1,N]$	$[N+2,2N]$	$a_{ij}(n-1,m-1,l-2N-1)$
$[N+2,2N]$	$[1,N]$	$[N+2,2N]$	$a_{ij}(n-2N-1,m-1,l-2N-1)$
$[1,N]$	$[N+2,2N]$	$[N+2,2N]$	$a_{ij}(n-1,m-2N-1,l-2N-1)$
$[N+2,2N]$	$[N+2,2N]$	$[N+2,2N]$	$a_{ij}(n-2N-1,m-2N-1,l-2N-1)$

Table D.1: Zero padding for three index kernel.

for $\gamma \in \{x,y,z\}$, and

$$A_{ij}(k_n,k_m,k_l) = \sum_{n,m,l=1}^{2N} A_{ij}(n,m,l) W_{2N}^{(k_n-1)(n-1)} W_{2N}^{(k_m-1)(m-1)} W_{2N}^{(k_l-1)(l-1)} \tag{D.39}$$

for $i,j \in \{1,2,3\}$. Next define the new sequences,

$$Y_x(k_n,k_m,k_l) = A_{11}S_x + A_{12}S_y + A_{13}S_z, \tag{D.40}$$

$$Y_y(k_n,k_m,k_l) = A_{21}S_x + A_{22}S_y + A_{23}S_z, \tag{D.41}$$

$$Y_z(k_n,k_m,k_l) = A_{31}S_x + A_{32}S_y + A_{33}S_z, \tag{D.42}$$

then the Cartesian components of the matrix vector product $\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}$ are given by the inverse transforms

$$\left[\bar{\bar{\mathbf{A}}} \cdot \bar{\mathbf{p}}\right]_\gamma = \frac{1}{(2N)^3} \sum_{k_n,k_m,k_l=1}^{2N} Y_\gamma(k_n,k_m,k_l) W_{2N}^{-(k_n-1)(n-1)} W_{2N}^{-(k_m-1)(m-1)} W_{2N}^{-(k_l-1)(l-1)}. \tag{D.43}$$