MONTE CARLO SOLUTION OF SCATTERING EQUATIONS FOR COMPUTER GRAPHICS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Matthew M. Pharr

June 2005

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Patrick M. Hanrahan Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Marc Levoy

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Ron Fedkiw

Approved for the University Committee on Graduate Studies.

Abstract

We present a mathematical framework for solving rendering problems in computer graphics that is based on scattering as its basic theoretical foundation, rather than the usual approach of simulating light transport and equilibrium. This framework can lead to more efficient solution methods than those based on previous methods as well as a number of new theoretical tools for solving rendering problems. We demonstrate the applicability of this approach to accurately and efficiently rendering subsurface scattering from geometric objects.

We first introduce a non-linear integral scattering equation that describes scattering from complex objects directly in terms of the composition of their lower-level scattering properties. This equation was first derived to solve scattering problems in astrophysics and has gone on to revolutionize approaches to transport problems in a number of fields. We derive this equation in a sufficiently general setting to be able to apply it to a variety of problems in graphics, which typically has problems with higher-dimensionality, more complexity, and less regularity than those in other fields. Methods to solve this equation have a divide-and-conquer flavor to them, in contrast to previous iterative methods based on the equation of transfer (i.e. the rendering equation). We apply Monte Carlo techniques to solve this scattering equation efficiently; to our knowledge, this is the first application of Monte Carlo to solving it in any field.

We next introduce Preisendorfer's *Interaction Principle*, which subsumes both scattering and light transport based approaches to transfer problems. It leads to a derivation of a set of *adding equations* that describe scattering from multiple objects in terms of how they

scatter light individually. We show how Monte Carlo techniques can be applied to solve these adding equations and apply them to the problem of rendering subsurface scattering.

Acknowledgements

I am extremely grateful to Pat Hanrahan for his support and encouragement over the years, starting from the process of my making my way to Stanford and then for giving me the opportunity to pursue a variety of topics in rendering over quite a few years as a part of the graphics lab. It has been an extraordinary experience from which I have learned an enormous amount; I feel extremely fortunate to have had these opportunities.

It has been a great privilege to be at the Stanford Graphics Laboratory and to work with and have many thought-provoking discussions with the great range of people who have passed through it over the years. Pat, Marc Levoy, Leo Guibas, and Ron Fedkiw have together made it into a magnificent working environment, filled with interesting people and great ideas. Thanks in particular to Eric Veach, Brian Smits, Steve Marschner, Olaf Hall-Holt, and Ravi Ramamoorthi for discussions and suggestions specifically regarding the work in this dissertation. In addition, Maneesh Agrawala, Joel Baxter, Andrew Beers, Ian Buck, Mike Cammarano, Milton Chen, Brian Curless, James Davis, Matthew "eeevil" Eldridge, Kayvon Fatahalian, Tim Foley, Brian Freyburger, Chase Garfinkle, Reid Gershbein, John Gerth, François Guimbretière, Chris Holt, Daniel Horn, Mike Houston, Greg Humphreys, Homan Igehy, Henrik Jensen, Craig "null" Kolb, Phil Lacroute, Bill Lorensen, Bill Mark, Ren Ng, John D. Owens, Hans Pedersen, Kekoa Proudfoot, Tim Purcell, Jonathan Ragan-Kelley, Szymon Rusinkiewicz, Pradeep Sen, Philipp Slusallek, Jeff "chipmap" Solomon, and Gordon Stoll, have all been great friends and colleagues along the way.

Thanks to Marc Levoy and Ron Fedkiw for also serving on my reading committee and to Art Owen for chairing the orals presentation. I am particularly grateful to Jim Arvo for flying up to serve on the orals committee and for a number of thoughtful comments on a draft of the dissertation.

I am thankful to Ada Glucksman and Heather Gentner for their help in navigating the university bureaucracy though this process and also to Kathi DiTommaso for her guidance as I applied for "the definitely positively very last candidacy extension" multiple times–this help was instrumental to this dissertation coming to completion!

The dragon model used in Chapter 4 was provided by the Stanford 3D Scanning Repository. Portions of this work appeared previously in Pharr and Hanrahan [PH00] and are reprinted with permission from the ACM. This work was supported in part by a Pixar Animation Studios graduate fellowship, DARPA contract DABT63-95-C-0085 and NSF contract CCR-9508579.

Finally, thanks to Deirdre, for everything.

Contents

List of Figures

Chapter 1

Introduction

Realistic image synthesis has long been a fundamental component of computer graphics. With applications ranging from entertainment (e.g. feature films, special effects, or video games) to lighting design and architecture, the need to realistically model the propagation and scattering of light in complex synthetic environments continues to challenge available computational resources.

1.1 The Rendering Problem

One part of the rendering process that has received much attention is the *global illumination* problem. It is based on the simulation of the inter-reflection of light in an environment in an effort to accurately describe the appearance of objects, accounting for all of the paths that light can take from light sources to objects in the scene. In recent years, global illumination has been placed on a solid theoretical and mathematical foundation. Important steps in this process include Kajiya's introduction of the rendering equation [Kaj86] and Arvo's work connecting the rendering equation with the more general equation of transfer [Arv93b, Arv95], thus solidifying connections with previous work in heat transfer, transport theory, and neutron transport. A distinguishing feature of these approaches is a focus on light

transport and equilibrium in the environment being modeled. In this thesis, we take a complementary approach to rendering, focusing on computing the scattering properties of the scene in order to describe how it reacts to an arbitrary incident illumination distribution. This approach leads to a variety of new theoretical machinery to apply to the rendering problem as well as a series of new efficient solution techniques.

The *modeling* problem is normally considered to be separate from the rendering problem. The set of surfaces that comprise the scene, the description of how they scatter light individually, and the description of any participating media at each point in the scene is usually considered to be given, such that the problems of modeling (e.g. how surfaces are represented mathematically) are usually considered to be mostly orthogonal to those of rendering.

One place where modeling and rendering do overlap is in computing *scattering functions* that describe how light scatters from surfaces or volumetric media. For example, the micro-scale structure of a surface may be modeled at a more detailed level than the global illumination algorithm can manage; for example, metallic surfaces are often modeled as a collection of randomly-oriented *microfacets* that reflect light like mirrors, or objects with layered structure, like skin or other biological objects, are often described as a collection of layers that individually scatter light in particular ways. In these sorts of situations, the scattering at the object's surface must either be approximated in some fashion, or a new rendering problem must first be solved at a finer scale in order to describe the object's scattering behavior at a higher level.

Previous work in rendering has largely focused on *light transport*; simulating the propagation of light from emitting sources to a sensor, such as an imaging system like a camera or the eye. In this thesis, we describe a new formulation of the general rendering problem that has generic *scattering* as its main focus, rather than light transport per se. The scattering function of an entire scene describes its generic response to illumination; if the scattering function of a scene can be efficiently evaluated, the scene can thus be rendered

directly. In general, the scattering function of a 3D scene is an eight-dimensional function and thus is too large to be pre-computed at a reasonable sampling rate. However, if the scattering function can be quickly evaluated for particular lighting and viewing directions, rendering is straightforward. This is the approach we focus on in this thesis.

We introduce new methods of computing the scattering behavior of complex objects and describe a general framework for computing aggregate scattering functions that directly describe the scattering from a collection of objects. This leads to a new theoretical framework for rendering problems in general. This kind of focus on scattering rather than light transport goes back over fifty years to work done in the field of astrophysics; it has since been applied to a variety of other areas. We apply it to a wider set of problems than it has been used for previously, make new connections between it and previous work in graphics, and develop new Monte Carlo methods for solving it that are applicable to a wider set of difficult problems than previous solution methods have been able to handle.

This work has a number of specific applications to problems in rendering. The most basic setting of interest is modeling *subsurface reflection* from layered objects. Given descriptions of how individual layers scatter light, we can compute solutions to the more general question of how light entering at the top or bottom of the stack of layers is reflected (or transmitted) from the object as a whole. In some cases, we can compute this scattering much more efficiently than previous approaches based on the rendering equation.

More generally, we derive a new method to compute scattering from volumetric media (as might be used to model a cloud, for example) in terms of scattering from sub-regions of the scene and the interactions between them. This approach can also lead to new methods for rendering scenes with general three-dimensional scattering geometry, new techniques for improved efficiency, and approximation methods that reduce computation time at some cost in accuracy. This approach can also lead to a new theoretical basis for *clustering*: approaches to rendering that try to increase efficiency by considering collections of objects as a single unit, rather than processing each of them individually. Along similar lines, the

level of detail problem involves choosing an appropriate representation for complex objects based on their visual contribution to an image; the generalized scattering framework that we describe applies to both of these directly.

1.2 Contributions

This thesis makes the following contributions:

- We introduce a pair of integral non-linear scattering equations to computer graphics that lead to new efficient methods for computing scattering from complex objects that are based on a divide-and-conquer approach rather than the iterative methods used for light transport problems.

- We develop the resulting scattering framework in a very general *ray space* setting. By deriving the theory in the most general three dimensional setting, we can easily apply it to a range of problems in graphics, all of which only require specialization of the ray space to a particular set of conditions.

- We apply the general solution technique of Monte Carlo sampling to solving the resulting integral equations. This method scales well to computing solutions to the equations in the general three dimensional setting, particularly in comparison to previous techniques which are intractable for the three-dimensional case.

- We introduce the adding equations to computer graphics; these describe scattering from objects in terms of scattering interactions between parts of them. This gives the correct theoretical foundation for computing scattering functions from the composition of objects.

- We also apply Monte Carlo to solving problems with the adding equations, which is a new solution approach for them as well. The generality of this solution technique

makes it possible to apply the adding equations to a wider variety of problems than they have been used for in other fields.

1.3 Thesis Organization

In the remainder of this thesis, we will first provide relevant background information about radiometry—a scientific (but phenomenological) description of light and its interactions with matter. This culminates with the introduction of the equation of transfer, an integral equation that describes light equilibrium in an environment with known scattering properties. We conclude with an introduction to scattering functions in a number of forms: they provide an important level of abstraction for the theory we will be developing.

We then derive a non-linear scattering equation in a general setting, which allows us to apply it to more complex scattering problems than it has been used for in other fields. We solve this scattering in integral (rather than differential) form, which allows us to apply Monte Carlo solution techniques to it. This solution method is compared to previous approaches, and we compare the efficiency of this approach to classic techniques from graphics based on the equation of transfer.

We next present Preisendorfer's interaction principle, which unifies the scattering and transport-based approaches to rendering and provides new tools for deriving new rendering algorithms. Furthermore, it solidifies the connection between these radiometric approaches and Maxwell's equations. We make connections between the interaction principle and previous rendering algorithms used in graphics and show its utility by deriving the adding equations, which describe overall scattering due to the composition of two objects with known scattering functions. They have been widely used previously in astrophysics and neutron transport and we show their utility for solving rendering problems in graphics. We describe new Monte Carlo methods for solving the adding equations and demonstrate their efficiency with a series of examples.

Finally, we conclude by summarizing this work and its contributions. We make additional connections with previous work in graphics and suggest a number of new directions for additional work.

Chapter 2

Background

We will first introduce background material necessary to the remainder of the chapters in this thesis. This material starts with a brief overview of concepts in radiometry, a system for describing and measuring light and electromagnetic radiation. We then describe basic concepts in radiative transfer, the study of how light interacts with and is scattered by participating media and surfaces in an environment. Finally, we introduce the general concept of scattering functions and show how computing values of scattering functions for particular lighting and viewing rays is the fundamental step for rendering imagery, regardless of the particular solution approach being taken.

2.1 Radiometry

Radiometry is a system of units and ideas for describing how light and power (or, more generally, electromagnetic radiation) are measured and how they propagate through an environment. See McCluney for a thorough introduction to the field [McC94]; our treatment in this section largely follows his development. Careful developments of the theory for application to computer graphics are available in Arvo [Arv95] and Veach [Vea97]. Alternatively, see Preisendorfer for an axiometric derivation of radiometric principles [Pre65,

Chapter 15].

Radiation is the propagation of energy through space. We generally are interested in the distribution of radiation over the electromagnetic spectrum; this is called the *spectral distribution*. We can also talk about *monochromatic radiation* at a particular wavelength λ. For applications in computer graphics, the portion of the electromagnetic spectrum from roughly 380mm to 800mm is generally what is of interest, as that corresponds to the set of visible wavelengths of light.

2.1.1 Basic radiometric quantities

There are four basic radiometric quantities relevant to rendering:

- Radiant flux, Φ, which has units of $[\text{watts}]$
- Irradiance, E $[\text{watts m}^{-2}]$
- Radiant intensity, I $[\text{watts sr}^{-1}]$
- Radiance, L $[\text{watts m}^{-2}\ \text{sr}^{-1}]$

Steradians, the solid angle unit, are denoted by sr. The *solid angle* subtended by a curve in space as viewed from some point x is equal to the area the curve covers when projected onto the unit sphere centered around x (see Figure 2.1). Thus, any sphere that encompasses x subtends a solid angle of 4πsr, since 4π is the area of the unit sphere. Similarly, the hemisphere subtends a solid angle of 2πsr.

Similarly, the *projected solid angle* subtended by a curve is equal to the area of the curve that results when the original curve is first projected onto the unit sphere around x and that curve is then projected to the unit disk around x; the area covered on the unit disk is the curves projected solid angle (Figure 2.2). The projected solid angle subtended by a sphere around x is 2π and the projected solid angle subtended by a hemisphere is π.

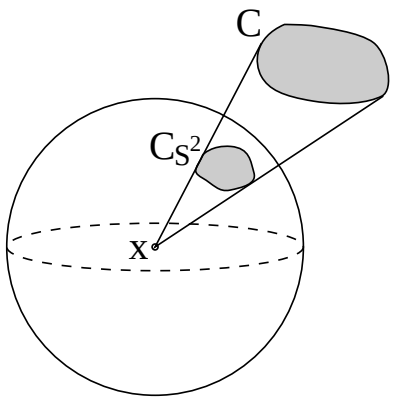


Figure 2.1: The solid angle subtended by a curve C in space with regard to a point x is equal to the area of the curve C_S^2 when projected onto the unit sphere surrounding x.

When writing integrals over sets of directions ω, we will denote the measure defined for the solid angle $d\mu(\omega)$ by $d\omega$. For the projected solid angle measure, we write $d\omega^\perp$ for $d\mu^\perp(\omega)$. Given surface normal $\mathbf{N}$, the Jacobian of the change of variables between the two measures is given by

$$\frac{d\omega^\perp}{d\omega} = |\omega \cdot \mathbf{N}|.$$

Radiant energy, Q, is energy passing through a given surface area over a given period of time. For the problems that we are interested in, the environment under consideration is in steady-state—i.e. the energy distribution has reached equilibrium and isn't changing over time (see Section 2.2). Therefore, radiant flux, Φ, or power, which is defined as the amount of energy passing through a surface per unit time, is a more generally useful quantity.

$$\Phi = \frac{dQ}{dt}$$

As Arvo points out, these kinds of radiometric definitions make the implicit assumption that light transport can be represented in a continuous formulation as is done by the above equation [Arv95, Kou69]. This assumption is necessary in that it allows us to use differential quantities and take their limits, but at some physical level, quantum physics means

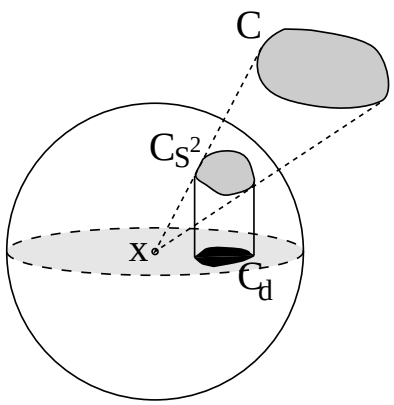


Figure 2.2: The projected solid angle is found by first projecting the curve C onto the sphere around x and then projecting this curve onto the unit disk around x that is perpendicular to the surface normal N.

that these limits can only be taken so far. Nonetheless, the accuracy of radiometry as a framework for predictive modeling of light transport is not affected.

Irradiance, E, is the flux per unit area on a surface at a given point x with surface normal N (see Figure 2.3). Note that a particular surface need not exist at x with normal N; it's just necessary that there be such a frame of reference in order to be able to define irradiance.

$$\mathrm{E} = \frac{\mathrm{d}\Phi}{\mathrm{dA}}$$

This is usually used to describe the incident distribution of illumination at a point on a surface, rather than a distribution of reflected light. Nevertheless, the concept is equally valid for both uses. (Sometimes the term "radiant exitance" is used to describe "irradiance leaving a surface" in order to differentiate it from irradiance arriving at a surface.) Irradiance is the basic quantity for describing radiation at a surface when the particular angular distribution is unimportant.

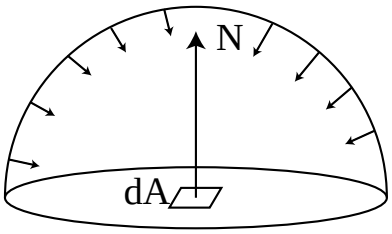


Figure 2.3: Irradiance is defined by the distribution of flux incident or exitent from a differential area dA on a surface with given surface normal.

By taking the limit of flux over directions rather than over areas, we find radiant intensity, I, flux per solid angle passing through a point in space.

$$\mathrm{I} = \frac{\mathrm{d}\Phi}{\mathrm{d}\omega}$$

Intensity is most useful for describing the angular distribution of emission from an infinitesimal point source; we will make only incidental use of this quantity in the remainder of this thesis.

Finally, radiance, L, is the solid angle and projected area density of flux, defined by

$$\mathrm{L} = \frac{\mathrm{d}^2\Phi}{\mathrm{d}\omega\,\mathrm{dA}^{\perp}},$$

where $\mathrm{dA}^{\perp}$ is the projected area measure on a plane perpendicular to ω. When we are considering radiance at a surface, we can equivalently write

$$\mathrm{L} = \frac{\mathrm{d}^2\Phi}{\mathrm{d}\omega\,\mathrm{dA}\,|\omega\cdot\mathbf{N}|} = \frac{\mathrm{d}^2\Phi}{\mathrm{d}\omega^{\perp}\,\mathrm{dA}}$$

Just as radiance can be expressed in terms of irradiance per differential solid angle, irradiance at a point can be computed from an incident radiance function:

$$\int_{\Omega} \mathrm{L}(x,\omega')\mathrm{d}\omega'^{\perp},$$

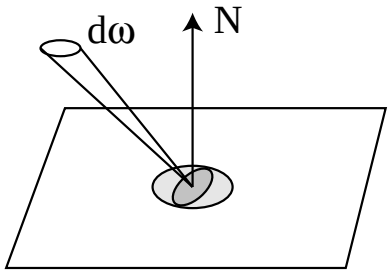


Figure 2.4: Radiance is the solid angle and projected area density of incident flux at a point x in a differential solid angle $d\omega$.

where Ω is the hemisphere above x and integration is with a projected solid angle measure.

Similarly, flux can be computed from a radiance function by

$$\Phi = \int_A \int_\Omega L(x, \omega') d\omega'^{\perp} dA,$$

and so forth.

2.1.2 Spectral issues

All of the quantities introduced thus far are spectral distributions that describe the distribution of the quantity over the electromagnetic spectrum. By taking another limit, we can define the *spectral* instances of each of them.

For example, spectral radiance at a given wavelength, L_λ, is just

$$L_\lambda = \frac{dL}{d\lambda}$$

As such, $L = \int_0^\infty L_\lambda(\lambda) d\lambda$.

A variety of representations for spectral distributions have been used in computer graphics; see e.g. Glassner [Gla95] for an overview. The two main approaches are to take point samples of a continuous spectrum at particular wavelengths (e.g. at the wavelengths corresponding to the red, green, and blue phosphors on a display), or to use a set of coefficients

of basis functions to represent the spectrum. For the remainder of this thesis, we will only consider monochromatic radiation and will therefore denote spectral radiance by L. For more general spectral representations, our methods can either be applied multiple times, once per sampled wavelength or basis function coefficient, or can be run once in parallel, computing results for all coefficients simultaneously (see also [Pre76, Section 1.8]).

A specialization of radiometry is *photometry*: in this setting, measured quantities are convolved with a standard response function $V(\lambda)$ that models humans' relative spectral response by wavelength. For example, luminance, the photometric equivalent of the radiometric quantity radiance, is equal to

$$683 \int_{380}^{770} V(\lambda) L_\lambda \, \mathrm{d}\lambda,$$

where λ is expressed in nanometers. Other radiometric quantities are defined similarly. Though photometry can be quite useful for practical applications where it is important to be able to quantitatively compare the perceived "brightness" of illumination, we will work entirely with radiometric quantities as they are more general and more fundamental.

2.1.3 The light field

An important operational concept is that of the *light field*. In a seminal paper from 1939, Gershun developed photometric concepts on a vector analysis foundation [Ger39]. In this work, we will make use of his development of the light field concept. In any environment (ignoring for now boundary conditions), radiance is defined at each point and in each direction, giving radiance distribution function (the light field), which is written as $\mathrm{L}(x, \omega)$.

2.2 Radiative Transfer and Geometrical Optics

Radiative transfer is the phenomenological study of the transfer of radiant energy through absorbing, scattering, and emitting media [Pre65]. Though it operates at the *geometrical optics* level, where macroscopic properties of light suffice to describe how light interacts with objects much larger than the wavelength of light, it is not at all uncommon to incorporate results from e.g. the wave optics level [HTSG91, Sta99]. These results just need to be expressed in the language of radiative transfer's basic abstractions.[1] At an even finer level of detail, quantum mechanics is needed to describe light's interaction with atoms; as direct simulation of quantum mechanical principles is unnecessary for solving rendering problems in computer graphics, the problem of the intractability of such an approach is avoided anyway.

In this geometrical optics framework, the assumption is made that light can be modeled as a flow of particles. Arvo has investigated this connection and made connections between previous work in *transport theory*, which applies classical physics to particles and their interactions to predict their overall behavior and global illumination algorithms [Arv93b, Arv95].

For the remainder of this thesis, we will assume that transport theory and geometrical optics are an adequate basis for the description of light and light scattering. As such, we will make following assumptions about the behavior of light:

- *Linearity*. The combined effect of two inputs to an optical system is always equal to the sum of the effects of each of the inputs individually.
- *Energy conservation*. More energy is never produced by a scattering event than there was to start with. This is in contrast to for example neutron transport problems where critical nuclear reactions can increase the amount of energy in the system.

[1] Preisendorfer has connected radiative transfer theory to Maxwell's classical equations describing electromagnetic fields [Pre65, Chapter 14]; his framework both demonstrates their equivalence and makes it easier to apply results from one world-view to the other. More recent work was done in this area by Fante [Fan81].

- *No polarization*. We will ignore polarization of the electromagnetic field; as such, the only property of light particles is their wavelength (or frequency). While most of the theory that we build this work upon has previously been extended to include the effects of polarization, we will ignore this effect for simplicity.

- *No fluorescence or phosphorescence*. In other words, we make the assumption that the behavior of light at one wavelength is completely independent of light's behavior at other wavelengths. As with polarization, it is relatively straightforward to include these effects in this work, but largely would serve to make the presentation more complex, with little practical advantage. More generally, we assume no "crossover" between wavelengths, which also rules out black body emission.

- *Steady state*. The environment has reached equilibrium, such that its radiance distribution isn't changing with time. This usually happens nearly instantaneously with light in realistic environments.

The most significant loss from assuming geometrical optics is that diffraction and interference effects cannot easily be accounted for. As noted by Preisendorfer, it is hard to fix this problem given these assumptions because, for example, in the presence of those effects the total flux over two areas isn't necessarily equal to sum of flux over each individually [Pre65, p. 24], thus breaking the linearity assumption.

2.3 Interactions

Given the radiative transfer and particle transport frameworks for reasoning about the behavior of light, we will now introduce the three main processes that affect the distribution of radiance in an environment. The first is absorption, which describes the reduction in radiance along a beam due to the absorption of energy (i.e. its conversion to another form of energy, such as heat) by a participating medium. Next is scattering, which describes how

light along a beam is scattered to different directions due to collisions with particles in the medium. Finally is emission, which describes the distribution of energy that is added to the environment from luminaires, etc. See Chandrasekhar, Preisendorfer, and Arvo for a more detailed introduction to these concepts [Cha60, Pre65, Arv93b].

All of the following properties may be *homogeneous* or *inhomogeneous*. Homogeneous properties are constant throughout the medium being considered, while inhomogeneous properties may vary arbitrarily throughout it.

2.3.1 Absorption

The first type of interaction that we will discuss is *absorption*. We start by defining the *beam transmittance* T_r, which is the fractional amount of radiance of a beam transmitted along a path from one point to another. It only accounts for radiance absorbed in the medium (sometimes called *true absorption*—radiance along other paths that is scattered *into* the beam that is being considered is handled separately). Given a beam starting at location x, going in direction ω and with length r, the *volume attenuation function* at a point x' along the beam is defined as

$$\mathrm{a}_r(x,\omega) = \frac{1 - T_r(x, x + r\omega)}{r},$$

where $T_r(x, x + r\omega)$ is the fraction of the original radiance at $x + r\omega$ that has been transmitted to x. In the limit, as $r \to 0$, we have the *absorption coefficient* at x in direction ω,

$$\sigma_a(x,\omega) = \lim_{r \to 0} \mathrm{a}_r(x,\omega),$$

which has units m^{-1}.

In most applications, σ_a is only a function of position and has the same value for all directions. Whichever form it takes will not make a difference for the remainder of this

discussion.

It can easily be shown that

$$\frac{\mathrm{d}T_r}{\mathrm{d}r} = -\sigma_a T_r,$$

from which it follows that

$$T_r = e^{-\int_0^r \sigma_a(x)\mathrm{d}x}.$$

For constant σ_a, *Beer's law*,

$$T_r = e^{-\sigma_a r},$$

follows directly. (See [Pre65, Section 17] for more detailed discussion of these ideas.)

2.3.2 Emission

Some *emission* processes are necessary for there to be any energy in an environment and thus for anything in it to be visible. Various chemical and thermal processes (or nuclear processes, e.g. in the case of the sun), convert energy into visible wavelengths of light which illuminate an environment. We will denoted emitted radiance at a point in a volume x, in a direction ω by $\mathrm{L_e}(x,\omega)$. We will assume that emission is a given property of the scene description and can be evaluated directly.

2.3.3 Scattering at a point

The third basic light interaction is *scattering*. As a beam of radiance propagates through a medium, it may collide with particles in the medium and be scattered into different directions (see Figure 2.5). This type of scattering is typically called *single scattering* to differentiate it from *multiple scattering*, which is the aggregate result of all radiance that goes through multiple single scattering events. A very closely related type of scattering is scattering at a surface; this will be described in Section 2.5.

Under the assumption that the individual particles that cause these scattering events

are separated by a few times the lengths of their radii, it is possible to ignore interactions between these particles when describing scattering at some location [van81]. The *phase function*, p, describes the angular distribution of scattered radiation at a point. A variety of phase functions that describe a variety of scattering media have been developed, ranging from parameterized models, which are used to fit a function with a small number of parameters to observed data, to analytic, which are derived by directly deriving the radiance distribution that results from scattering from, for example, spherical particles.

Phase functions are typically defined so that they are normalized—for all directions ω,

$$\frac{1}{4\pi}\int_{S^2} p(\omega \to \omega')\, d\omega' = 1. \tag{2.1}$$

In most naturally-occurring media, the phase function is a function of the angle between the two directions ω and ω'; such media are called isotropic and these phase functions are often written as $p(\cos\theta)$. In an anisotropic medium, such as one with crystalline-type structure, the phase function is a function of the values of each of the two angles. An important property of naturally-occurring phase functions is that they are *reciprocal*: the two directions can be interchanged and the phase function's value remains unchanged.

Phase functions can be isotropic or anisotropic as well. The isotropic phase function is independent of either of the two angles and always has a value of 1 (in order to satisfy Equation 2.1. An anisotropic phase function depends on either the angle between the two directions or the two directions themselves, depending on if the medium is isotropic or anisotropic, respectively.

Phase functions are often described by an asymmetry parameter, g, that is the average value of the product of the phase function with the cosine of the angle between ω' and ω. The range of g is from -1 to 1, corresponding to total back-scattering to total forward

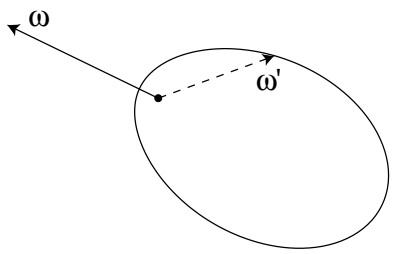


Figure 2.5: The phase function describes the distribution of scattered radiance in directions ω' at a point, given incident radiance along the direction ω. Here we have plotted the Henyey-Greenstein phase function with an asymmetry parameter g equal to 0.5.

scattering, respectively. For any ω,

$$g = \frac{1}{2}\int_{S^2} \mathrm{p}(\omega \to \omega')(\omega \cdot \omega')\mathrm{d}\omega' \tag{2.2}$$

Thus, isotropic scattering corresponds to a g of zero. Any number of phase functions can satisfy Equation 2.2; the g value alone is not enough to uniquely describe a scattering distribution. Nevertheless, the convenience of being able to easily convert a complex scattering distribution into a simple parameterized model often outweighs the loss in accuracy. Our examples in Chapter 3.4 are based on a phase function due to Henyey and Greenstein [HG41]; it has been widely used and is both computationally efficient and fits many observed phase functions well.[2]

$$\mathrm{p}(\omega \to \omega') = \frac{1-g^2}{(1+g^2-2g(-\omega \cdot \omega'))^{3/2}} \tag{2.3}$$

However, none of the theory that we develop depends on the particular expression of a phase function being used. More complex phase functions that aren't described well with a single asymmetry parameter are often modeled with a weighted sum of phase functions

[2]This equation is slightly different than many other authors' since we use the convention that both directions are pointing outward from the point where scattering happens.

like Henyey-Greenstein with different parameter values:

$$p(\omega \rightarrow \omega') = \sum_{i=1}^{n} w_i \, p_i(\omega \rightarrow \omega') \tag{2.4}$$

where the weights, w_i necessarily sum to one so that the normalization condition, Equation 2.1, holds.

Detailed discussion of scattering and phase functions and derivations of phase functions that describe scattering from independent spheres, cylinders, and other simple shapes can be found in van de Hulst [van81]. In particular, extensive discussion of the commonly-used Mie and Rayleigh scattering models (which describe scattering from particles approximately the size of or larger than the wavelength of incident radiation and particles much smaller than the wavelength of incident radiation, respectively) is available there. Hansen and Travis's survey article is also a good introduction to the variety of commonly-used phase functions [HT74].

2.3.4 Scattering coefficient

In addition to the the phase function, we need to know how often scattering events take place. The scattering coefficient σ_s describes the probability of a scattering event per unit distance traveled in the medium. When added to the absorption coefficient, we have the *attenuation coefficient*, or *interaction coefficient*,

$$\sigma_t = \sigma_a + \sigma_s,$$

which gives the total reduction in energy due to out-scattering and absorption.

A useful quantity based on the absorption coefficient is the *optical thickness* between two points. It is just the integral over the line between the two points of the absorption

coefficient at each point.

$$\int_x^{x'} \sigma_t(x'')\mathrm{d}x''$$

The *albedo* is a useful quantity based on the scattering and attenuation coefficients; it gives the fraction of the incident radiance that is re-scattered at a scattering event:

$$\alpha = \frac{\sigma_s}{\sigma_t}.$$

High-albedo media, with α close to one, re-scatter most of the incident light at each scattering interaction, while low-albedo media, with α close to zero, absorb most of it.

2.4 The Equation of Transfer

The *equation of transfer* is the fundamental equation that governs the behavior of light in some medium that absorbs, emits, and scatters radiation [Cha60]. It describes the equilibrium distribution of radiance in terms of the incident radiance in the medium and the medium's scattering properties. Most rendering algorithms in computer graphics focus on solving the equation of transfer in order to compute the radiance leaving the scattering medium or geometric objects and arriving at a sensor (e.g. the virtual film plane of a camera model.)

As radiance travels along a beam, a number of processes contribute to change its distribution. Radiance can be increased due to emission and *in-scattering*, radiance along other beams that is scattered into the path of the beam under consideration. Conversely, radiance can be decreased due to absorption and *out-scattering*, radiance that is scattered into other beams.

The equation of transfer describes this process. In its most basic form, it is an integro-differential equation that describes how the radiance along a beam changes at a point. It can be derived by subtracting the effects of the scattering processes that reduce energy along

the beam (absorption and out-scattering) from the processes that increase energy along the beam (emission and in-scattering). Here we will assume that the medium has a constant index of refraction—i.e. a beam follows a straight line path; see Preisendorfer [Pre65, Section 21] for the derivation in the more general setting.

The *source function* is the amount of new light at a point in a direction due to emission and in-scattered light from other points in the medium:

$$\mathrm{J}(x,\omega) = \mathrm{L_e}(x,\omega) + \frac{1}{4\pi}\int_{\mathcal{S}^2} \mathrm{p}(x,\omega' \to \omega)\mathrm{L_i}(x,\omega')\mathrm{d}\omega'$$

Consider now a differential volume along a beam of radiation. The beam is parameterized along its direction by a variable $t \geq 0$ such that points on the beam are given by $x + t\omega$. Now by combining the source function with an expression for the loss in radiation due to attenuation and out-scattering, we have the integro-differential form of the equation of transfer [Cha60, Arv93b]:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{L}(x') = -\sigma_t \mathrm{L}(x') + \mathrm{J}(x') \tag{2.5}$$

With suitable boundary conditions, this can be transformed to a purely integral equation. If we assume that there are no surfaces in the scene, we have:

$$\mathrm{L}(x,\omega) = \int_0^\infty e^{-\int_0^{t'} \sigma_t(x'')\mathrm{d}t''} \mathrm{J}(x',\omega)\mathrm{d}t', \tag{2.6}$$

where $x' = x + t'\omega$ and similarly for x'' and t''.

More generally, if there are reflecting and/or emitting surfaces in the scene, we have:

$$\mathrm{L}(x,\omega) = e^{-\int_0^t \sigma_t(x')\mathrm{d}t'} \mathrm{L_o}(x_0,-\omega) + \int_0^t e^{-\int_0^{t'} \sigma_t(x'')\mathrm{d}t''} \mathrm{J}(x',\omega)\mathrm{d}t' \tag{2.7}$$

where t is the distance along the ray to the first surface, and $\mathrm{L_o}$is the radiance leaving the

surface, accounting for both emission and reflected incident radiance (see Equation 2.8.)

The equation of transfer was first introduced to graphics by Kajiya and von Herzen [KV84]; Rushmeier was the first to compute solutions of it in a general setting [Rus88]. Arvo was the first to make the essential connections between previous formalizations of light transport in graphics and the equation of transfer and radiative transfer in general [Arv93b].

2.4.1 The rendering equation

As explained by Arvo [Arv93b], Kajiya's *rendering equation* [Kaj86] is a particular instance of the equation of transfer with particular boundary conditions. However, we will generally distinguish between the two equations, using the term "rendering equation" to describe the equation of transfer applied to rendering scattering from surfaces.[3]

The rendering equation says that, assuming all surfaces are opaque, the radiance leaving a surface from a point in a given direction is given by the sum of the emitted radiance (if the surface is emissive) and the incident radiance scattered by the surface,

$$\mathrm{L_o}(x,\omega) = \mathrm{L_e}(x,\omega) + \int_\Omega f_\mathrm{r}(\omega' \rightarrow \omega)\mathrm{L_i}(x,\omega')\mathrm{d}\omega'^{\perp}, \tag{2.8}$$

where f_r is the bidirectional reflectance distribution function; it describes scattering from surfaces in a manner similar to the phase function and is explained in detail in Section 2.5. Here we have written the integral in the rendering equation over incident radiance over the hemisphere at x; Kajiya's original formulation was an integral over all surfaces in the scene. As the two are equivalent, we used this expression in order to match our nomenclature for the equation of transfer.

[3] Authors in graphics have often used the term *volume rendering equation* to describe the equation of transfer [Kru90].

The second term of the rendering equation is often known as the the *reflectance equation* [CW93, Chapter 2]; it describes the reflected radiance distribution at a surface due to a given distribution of incident radiance $\mathrm{L_i}$.

$$\mathrm{L_o}(x,\omega) = \int_\Omega f_\mathrm{r}(\omega' \rightarrow \omega)\mathrm{L_i}(x,\omega')\mathrm{d}\omega'^{\perp} \tag{2.9}$$

2.5 General Scattering Functions

While most previous work on rendering algorithms for computer graphics has focused on solving the equation of transfer, typically by simulating the propagation of incident illumination through the environment, this thesis is based on a different approach, computing the scattering function of the scene being rendered, which describes how an incident radiance function that represents the illumination of the scene is transformed into an outgoing radiance function describing the distribution of illumination leaving the scene.

Consider a scene to be rendered inside a convex volume illuminated by an incident radiance function $\mathrm{L_i}$. The radiance leaving the scene at some point x on the boundary of the volume in a direction ω can be written in terms of the integral of incident radiance over the volume

$$\mathrm{L_o}(x,\omega) = \int_A \int_\Omega \mathrm{S}(x_i,\omega_i \rightarrow x_o,\omega_o)\mathrm{L_i}(x_i,\omega_i)\,\mathrm{d}x_i\mathrm{d}\omega_i.$$

The scattering function, S, characterizes how incident radiance is transformed into exitant radiance by the scene. Under the radiometry assumptions, it fully characterizes the light scattering properties of the scene. In its most general form, as described here, it is an eight dimensional function, which indicates that precomputing it or representing it in a tabularized form is likely to be impractical in general. However, if more efficient algorithms can be found to evaluate S for given pairs of incident and exitant rays than using the equation of transfer to compute exitant radiance, the scattering function-based approach can be a

preferable approach for rendering. General scattering functions of this form were introduced to computer graphics simultaneously by Pharr and Hanrahan [PH00] and Debevec et al [DHT+00].

Just as the phase functions defined in Section 2.3.3 abstract away low-level particle scattering processes at the wave optics level, general scattering functions describe the scattering behavior of a surface or general object. Such scattering functions are similarly important abstractions in that they describe scattering behavior phenomenologically such that lower-level scattering processes are hidden. The remainder of this section covers the background necessary to show how being able to efficiently evaluate the scattering function for objects is essentially the goal of physically-based rendering algorithms.

2.5.1 Previous Work: Scattering from Surfaces

Limited cases of general scattering functions have been widely used in graphics to describe scattering from surfaces. Most of this work is based on the framework of Nicodemus et al., who introduced the bidirectional subsurface reflectance distribution function (BSSRDF) which describes light reflection from surfaces in a very general fashion [NRH+77]. Given a beam of incident radiation $d\Phi$ from direction ω_i that irradiates a differential surface area dA_i at point x_i, the amount of reflected radiance at point x_o in direction ω_o due to $d\Phi$ is denoted by dL_o. Due to the linearity of light reflection, $d\Phi$ and dL_o are proportional.

This factor of proportionality was defined to be the BSSRDF. Denoting it by f_{ss}, we have a function of two positions on the surface and two directions, $f_{ss}(x_i, \omega_i \rightarrow x_o, \omega_o)$ (see Figure 2.6). This is a scattering function that accounts for inhomogeneity in all dimensions underneath the surface; it gives reflectance along an outgoing ray due to illumination along an incoming ray.

If we have a surface with a given BSSRDF that is illuminated by a given incident radiance distribution, rendering an image of that surface requires that we compute the radiance

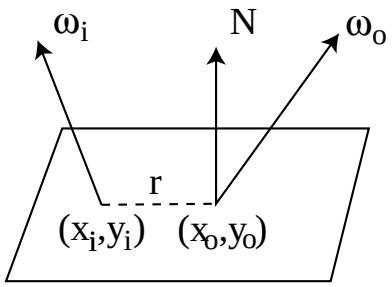


Figure 2.6: The BSSRDF at a point on a surface (x_o, y_o) is given by the ratio of reflected radiance in the direction ω_o to incident radiation $d\Phi$ at the point (x_i, y_i) from the direction ω_i.

leaving it at a set of points x in directions ω. For a BSSRDF, the integral that must be evaluated for each such (x, ω) is

$$L_o(x, \omega) = \int_A \int_\Omega f_{ss}(x_i, \omega_i \rightarrow x_o, \omega_o) L_i(x_i, \omega_i) \, dx_i d\omega^\perp.$$

In other words, we must consider the hemispherical distribution of incident radiance at all points x_i on the surface's boundary and compute the product of the radiance with the BSSRDF's value.

In practice, this integral is approximated with a numeric integration technique (e.g. Monte Carlo integration). The L_i term can usually be evaluated efficiently, and most of the computational expense is in computing $f_{ss}(x_i, \omega_i \rightarrow x_o, \omega_o)$ for many pairs of points and directions.

An often-useful set of simplifying assumptions transforms the BSSRDF to the BRDF (bidirectional reflectance distribution function). The BRDF has been widely used in graphics as the basic abstraction that describes scattering at a surface (e.g. [CW93, Chapter 2]). Given the simplifying assumption that the surface is homogeneous and uniformly illuminated over a reasonably large area (or, equivalently, that light exits the surface at the same point it enters), the BSSRDF doesn't depend on the location of x_o, but only on the distance between x_i and x_o. Writing this distance as r, we can reduce the reflection function to a

four-dimensional function over pairs of angles [NRH+77]:

$$f_r(\omega_i \rightarrow \omega_o) = \int_{A_i} f_{ss}(r; \omega_i \rightarrow \omega_o) dA_i$$

and

$$f_r(\omega' \rightarrow \omega) = \frac{dL_o(\omega)}{dE(\omega')} = \frac{L_o(\omega)}{dL_i(\omega') d\omega'^{\perp}}. \quad (2.10)$$

Given the BRDF, the radiance leaving a point on a surface in a given direction only requires an integral over the hemisphere, as given by the reflection equation, 2.9.

2.5.2 Formal Inversion of the Equation of Transfer

One approach to computing values of scattering functions is via the formal solution of the inverse of the light transport equation [Pre65, Section 22]. Indeed, previous rendering algorithms used in graphics can be understood to effectively be based on computing the values of scattering functions. Veach and Guibas derived rendering algorithms based on recursive expansion of this solution operator [VG94, Vea97] and Lafortune used the Neumann expansion of the solution operator to derive recursively-defined integral equations that describe scattering from a collection of surfaces; he called this the global reflectance distribution function (GRDF) and also used it to derive new light transport algorithms [LW94, Laf96]. Arvo and Veach have extensively investigated properties of the solution operator and conditions for its existence[Arv95, Vea97].

In order to simplify equations for future discussions, we will introduce a set of linear operators (see e.g. Taylor and Lay for an introduction to operators on vector spaces[TL80]) that describe scattering and propagation of radiance in an environment based on a notation developed by Preisendorfer [Pre65]. Similar frameworks focused on the rendering equation, rather than the equation of transfer, have been developed by Arvo [Arv93a, Arv95] and Veach [Vea97].

Consider a convex optical medium with no emitting objects inside it that is illuminated by an incident radiance distribution, L_b, at its boundary. At each point and in each direction inside the medium, we can define a radiance distribution due to this direct illumination by accounting for the attenuation from the boundary to the point; we will denote this by $\mathrm{L}_0(x, \omega)$.

Now given the radiance field L_0, we can define the *first order radiance field* L_1, which gives radiance at a point x in direction ω due to incident radiance that has scattered one time. At each point x' along the line from x to the boundary in the direction w, scattered direct illumination is found by the integral over the sphere; this radiance is then attenuated along the line from x' to x.

$$\mathrm{L}_1(x,\omega) = \int_0^t \int_{\mathcal{S}^2} \mathrm{L}_0(x',\omega')\mathrm{p}(\omega' \to \omega)\mathrm{d}\omega' T_r(x',\omega)\mathrm{d}t'$$

where r is the distance from x to the boundary of the medium and where $x' = x + t'\omega$.

L_1 can be defined as a two stage process. First we define a phase scattering operator $\mathbf{P}$ by

$$(\mathbf{P}f)(x,\omega) = \int_{\mathcal{S}^2} \mathrm{p}(\omega' \to \omega)\, f(x,\omega')\mathrm{d}\omega'.$$

Next we define a transmission operator $\mathbf{T}$, by

$$(\mathbf{T}f)(x,\omega) = \int_0^r T_r(x, x + r\omega)\, f(x,\omega)\, \mathrm{d}t'$$

We can then define in general the *j-th order radiance function* by

$$\mathrm{L}_j = \mathbf{TP}\mathrm{L}_{j-1}.$$

This expression gives the radiance function for light that has been scattered exactly j times after emission.

If we combine **T** and **P** into a single *light transport operator*, $\mathbf{K} = \mathbf{TP}$, then the light field in the medium accounting for all orders of scattering is

$$\mathrm{L} = \sum_{i=0}^{\infty} \mathrm{L}_i = \sum_{i=0}^{\infty} \mathbf{K}^i \mathrm{L}_0, \tag{2.11}$$

where $\mathbf{K}^i$ denotes i applications of **K** and $\mathbf{K}^0$ is the *identity operator* **I**. This infinite series, which is known as the *natural solution* of the equation of transfer, formally satisfies the equation of transfer and converges under mild requirements of physically correct scattering behavior.

We can write the formal inversion of the operator equation, 2.11 [Pre65],

$$\mathrm{L} = (\mathbf{I} - \mathbf{K})^{-1} \mathrm{L}_0 \tag{2.12}$$

Computing this *solution operator*, $(\mathbf{I} - \mathbf{K})^{-1}$ is the implicit goal of many applications in rendering. For example, if an incident radiance distribution L_0 is known (e.g. from a set of light sources in a synthetic scene), and if the object to be rendered by described by given attenuation coefficient functions, phase functions, etc., in order to render an image of the object, it's necessary to compute the radiance leaving the object, L. Doing so requires computing the value of the solution operator for the incident radiance distribution and the appropriate outgoing directions for the image. More generally, the solution operator describes the response of an object to any incident radiance distribution.

A common solution method has been to expand the solution operator into its Neumann series; this is the approach taken by Kajiya, for example [Kaj86].

$$(\mathbf{I} - \mathbf{K})^{-1} = \mathbf{I} + \mathbf{K} + \mathbf{K}^2 + \cdots$$

See Arvo [Arv93a] for an overview of integral equation solution techniques and their application to the equation of transfer and the rendering equation.

2.5.3 Previous Work in Graphics

Computing scattering functions that hide the complexity of light scattering from surfaces has long been a research problem in graphics and optics. In graphics, this work has been done in the context of reflection from surfaces modeled by the BRDF and the BSSRDF as well as scattering functions that model scattering from complex geometry.

BRDF Models

Representative examples of BRDF scattering models in grapihcs include the Torrance–Sparrow reflection model [TS67], an analytic approximation to light scattering from rough surfaces; Blinn's model for dusty surfaces, which uses a single-scattering approximation [Bli82]; Kajiya's discussion of replacing complex geometry with reflection functions [Kaj85]; and Westin et al.'s computation of BRDF samples by simulating light scattering from micro-geometry [WAT92].

When no closed-form expression or approximation for multiple scattering at a surface is available, previous work has either ignored multiple scattering (e.g. [Bli82]), or based solutions on the equation of transfer and the definition of the BRDF (e.g. [WAT92]), where reflected radiance in the outgoing direction is computed given differential irradiance from the incident direction.

BSSRDF Models

Miller and Mondesir computed hypersprites that encoded specular reflection and refraction from objects [MM98], and Zongker et al. have described an apparatus for computing the scattering and transmission functions of glossy and specular real-world objects [ZWCS99].

Dorsey et al. have rendered rich images of stone and marble by computing BSSRDFs at rendering time [DEL$^+$99]. Their solutions are based on the equation of transfer and photon mapping to accelerate multiple scattering computations, and they clearly showed

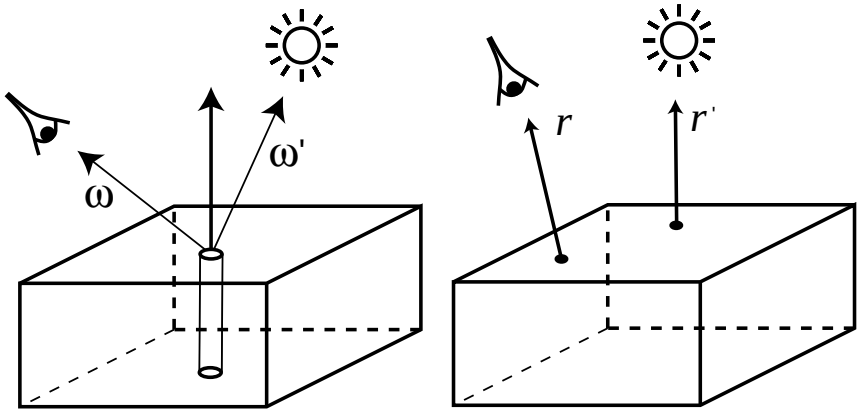


Figure 2.7: Basic viewing geometry for the 1D (left) and 3D (right) scattering functions. All vectors and rays are specified in the outgoing direction.

the importance of this effect for some materials.

More recently, Jensen et al have rendered subsurface scattering from BSSRDFs using a dipole method that approximates multiple scattering very efficiently under a number of simplifying assumptions [JMLH01, JB02]. This approach is particularly effective in the high albedo case, where light is scattered many times before being absorbed. However, this method is derived under the assumptions that the medium is semi-infinite and homogeneous; while they have shown the applicability of the method in more general settings, the error introduced by doing so is difficult to quantify.

Scattering from Complex Objects

In recent years, a number of researchers have worked on computing scattering functions that describe the aggregate scattering behavior of complex volumetric and geometric objects. Kajiya and Kay's volume texels were an early example [KK89], and Neyret extended their framework to include more general geometries and demonstrated applications to reducing aliasing due to level-of-detail changes [Ney98]. Rushmeier et al. approximated scattering from clusters of geometry by averaging the reflectance of surfaces hit by random

rays [RPV93]. Sillion and Drettakis approximated occlusion due to complex objects as volume attenuation functions [SD95] and Sillion et al. approximated aggregate scattering functions from clusters of objects [SDS95]. None of these approaches accounts for multiple scattering inside the object or for light that enters the object at a different point than it exits as general scattering functions do. Debevec et al. measured reflection functions of human faces using a general scattering approach [DHT$^+$00].

Chapter 3

Nonlinear Integral Scattering Equations

This chapter introduces a pair of integral scattering equations that express the scattering function of participating medium in terms of the scattering functions of smaller subsets of it. In contrast to computing the scattering function via inversion of the equation of transfer, where many terms of the $(\mathbf{I}-\mathbf{K})^{-1}$ expansion must be evaluated for accurate results if the scattering albedo is high, the integral scattering equations are integrals over depth through the medium. As such, they can be evaluated more efficiently than the equation of transfer. Furthermore, because they include terms that themselves have scattering functions, there are opportunities of approximate these terms in a manner that isn't possible with previous approaches, leading to accurate solutions more efficiently.

3.1 Ray Space and Operator Notation

Before deriving these scattering equations, we will first introduce the notation we will use throughout the remainder of this chapter.

Previous work in graphics has used a variety of parameterizations of surfaces and directions for the expression of the rendering equation (e.g. Kajiya used an integral over pairs of points on surfaces). Veach has introduced abstractions based on *ray space* that have a

x	Generic point
ω	Generic direction
r	A ray through space, with origin $x(r)$ and direction $\omega(r)$
μ_r	Cosine of ray's direction with surface normal
$\delta(x)$	Delta function: Kronecker or Dirac, depending on context
$\mathcal{S}^2$	The sphere of all directions
Ω	The hemisphere around the $+z$ direction
$\mathcal{M}^2$	A 2D manifold
$\mathcal{R}$	Ray space: a set of rays going through a set of locations in a set of directions
$\mathcal{R}_{\mathcal{M}^2}$	The set of rays originating on a two dimensional manifold $\mathcal{M}^2$
$\mathrm{L}(r)$	Radiance along the ray r
$\mathrm{p}(x, \omega' \to \omega)$	Phase function at a point.
$\mathrm{k}(r' \to r)$	Scattering kernel
$\mathrm{S}(r' \to r)$	Scattering function for light reflected along ray r due to incident light along ray r'
$\sigma_a(x)$	Volume absorption coefficient at x
$\sigma_s(x)$	Volume scattering coefficient
$\sigma_t(x)$	Volume attenuation coefficient, $\sigma_s(x) + \sigma_a(x)$
$\alpha(x)$	Albedo $\sigma_s(x)/\sigma_t(x)$
z	Depth in one-dimensional medium
$\mathrm{R}(z, \omega_i \to \omega_o)$	Reflection function from slab of depth z
$\mathrm{T}(z, \omega_i \to \omega_o)$	Transmission function from slab of depth z

Figure 3.1: Table of notation.

number of advantages: in addition to simplifying and clarifying formulas, ray space makes clear that any particular parameterization of surfaces and directions is an arbitrary choice, mathematically equivalent to any other [VG95, Vea97].

In this setting, ray space $\mathcal{R}$ is the set of rays given by the Cartesian product of points in three-space $\mathbb{R}^3$ and all directions $\mathcal{S}^2$: $\mathcal{R} = \mathbb{R}^3 \times \mathcal{S}^2$. We will define two specializations of $\mathcal{R}$. First is $\mathcal{R}_{\mathcal{M}^2}$, which is the subset of $\mathcal{R}$ where all rays start on a given two-dimensional manifold $\mathcal{M}^2$: $\mathcal{R}_{\mathcal{M}^2} = \mathcal{M}^2 \times \mathcal{S}^2$. A particular instance of $\mathcal{R}_{\mathcal{M}^2}$ that is often useful is

$\mathcal{R}_{\mathcal{M}^2(z')}$, where the manifold is the plane at $z = z'$. Another useful specialization is to limit the directions of rays $\mathcal{R}$ to the hemisphere around the surface normal of the manifold at the ray's origin; we denote this by $\mathcal{R}^+_{\mathcal{M}^2}$. The negation of a ray $-r$ is defined as the ray with the same origin as r but going in the opposite direction.

The scattering kernel k describes light scattering at a point. In ray space it is defined as

$$\mathrm{k}(r' \to r) = \delta(x(r) - x(r'))\ \sigma_s(x(r))\ \mathrm{p}(x(r), \omega(r') \to \omega(r)),$$

where $\mathrm{p}(x(r), \omega(r') \to \omega(r))$ is the phase function at the point $x(r)$ for scattering from $\omega(r')$ to $\omega(r)$ and we have included the scattering coefficient σ_s in k in order to simplify subsequent formulas. Note that it is zero unless the two rays share the same origin.

In contrast to the phase function and scattering kernel, the scattering function $\mathrm{S}(r' \to r)$ is potentially non-zero for any pair of rays because of multiple scattering; it is not necessary that the rays meet at a point for light along one ray to affect the response along another. Though the general scattering function is ten-dimensional (since it is based on two rays, each of which require five dimensions), when we are considering scattering from a specific object, it is often more convenient to consider the eight-dimensional specialization where all rays originate on a parameterized two-dimensional manifold that bounds it.

In order to be able to do integrals over $\mathcal{R}_{\mathcal{M}^2}$ and $\mathcal{R}_{\mathcal{M}^2(z)}$, we define a differential measure:

$$\mathrm{d}r = \mathrm{d}\omega(\omega(r))\,\mathrm{dA}^{\perp}(x(r)) = \mu_r\,\mathrm{d}\omega(\omega(r))\,\mathrm{dA}(x(r))$$

where $x(r)$ is the origin of r, $\omega(r)$ is its direction, A is the area measure on $\mathcal{R}_{\mathcal{M}^2}$, and $\mathrm{d}\omega$ is the differential solid angle measure.

Given an object's scattering function, outgoing radiance along a ray r is computed by

integrating its product with incident radiance over the object's boundary.

$$\mathrm{L_o}(r) = \frac{1}{4\pi} \int_{\mathcal{R}^+_{\mathcal{M}^2}} \frac{\mathrm{S}(r' \to r)}{\mu_r \mu_{r'}} \mathrm{L_i}(r') \, \mathrm{d}r'$$

This is the three-dimensional analog to the reflectance equation, 2.9, which integrates the product of incident radiance and the BRDF at a point to compute outgoing radiance. Its added complexity stems from the fact that incident light scatters inside the object and may exit far from where it entered.

We will define operators **k** and **S**, where bold text signifies the operator and Roman text its kernel. Both operators are defined such that applying them to other functions gives:

$$(\mathbf{S}\, f)(r) = \int_{\mathcal{R}^+_{\mathcal{M}^2}} \frac{\mathrm{S}(r' \to r)}{\mu_r \mu_{r'}} f(r') \, \mathrm{d}r'$$

We can define compositions like **kS**, or $\mathbf{S}_a\mathbf{S}_b\mathbf{S}_c$, etc. These will be useful in computing new scattering functions that describe the scattering of multiple objects in terms of their individual scattering functions (see Chapter 4).

$$(\mathbf{S}_1 \ldots \mathrm{S}_n)(r' \to r) = \int_{\mathcal{R}^+_{\mathcal{M}^2}} \cdots \int_{\mathcal{R}^+_{\mathcal{M}^2}} S_n(r' \to r_1) \cdots S_1(r_{n-1} \to r) \frac{\mathrm{d}r_{n-1}}{\mu^2_{r_{n-1}}} \cdots \frac{\mathrm{d}r_1}{\mu^2_{r_1}} \qquad (3.1)$$

3.2 Derivation of the Scattering Equation

With operator notation in hand, we will derive a pair of general integro-differential scattering equations in ray space. These equations describe how the scattering function of an object changes as layers with known scattering properties are added or removed from it. It can either be solved in integro-differential form or as a purely integral equation. Our derivation follows the *invariant imbedding* method [BK56, BKW60, Pre58b, BKP63, BW75].

We will consider scattering from objects in an axis-aligned rectangular region of space

with height z. This does not require that the object be parallelepiped-shaped; it is just a convenient parameterization of space. This parameterization also makes it possible to ignore the issue of non-convex regions of space, where illumination may exit and later re-enter the space. That setting is also tractable, though the notation is more complex.

3.2.1 Reflection

The first scattering equation describes reflection from the medium, where here reflection means that both rays originate on either the top or bottom of the medium and point in the same hemisphere.

We will first consider the change in scattering behavior of this object as thin layers Δz are added on top of it. Multiple scattering in Δz occurs with probability $o(\Delta z^2)$, and so we will gather all multiple scattering in an $o(\Delta z^2)$ term. Later in the derivation we will divide by Δz and take the limit as $\Delta z \to 0$, at which point all of the $o(\Delta z^2)$ terms disappear. There are five remaining types of scattering events that need to be accounted for that are not in the $o(\Delta z^2)$ term (see Figure 3.2):

1. **S**: Light that is attenuated in Δz, scattered by the original object, and attenuated again in Δz.
2. **k**: Light that is scattered in Δz so that it leaves the new layer without reaching the original object.
3. **kS**: Light that is scattered in Δz so that it passes into the original object, is scattered, and then is attenuated in Δz as it exits.
4. **Sk**: Light that is attenuated in Δz, scattered by the object, and is then scattered by Δz such that it leaves the object.
5. **SkS**: Light that is attenuated in Δz, scattered by the object, scattered in Δz back into the object, scattered again by the object, and then attenuated again.

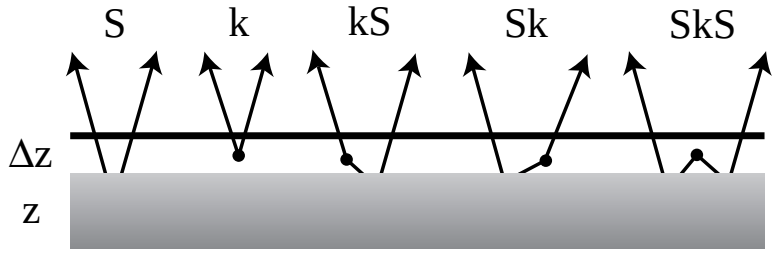


Figure 3.2: The five types of scattering events to be considered in the invariant imbedding derivation of the scattering equation. The *S* events reflect the aggregate multiple scattering inside the *z* slab. All other scattering events, such as kSk, are gathered in an $o(\Delta z^2)$ term in Equation 3.2.

Accounting for each of the five modes of scattering in turn, a new scattering function for $z+\Delta z$ can be written

$$\mathbf{S}(z+\Delta z) = \mathrm{e}^{-\sigma_t(\Delta z/\mu_i)}\mathrm{e}^{-\sigma_t(\Delta z/\mu_o)}\Big(\mathbf{S}(z)+ \mathbf{S}(z,z+\Delta z)+\mathbf{S}(z,z+\Delta z)\mathbf{S}(z)+\mathbf{S}(z)\mathbf{S}(z,z+\Delta z)+ \mathbf{S}(z)\mathbf{S}(z,z+\Delta z)\mathbf{S}(z)+o(\Delta z^2)\Big), \quad (3.2)$$

where $\mathbf{S}(z)$ is the scattering operator for a slab of thickness z and $\mathbf{S}(a,b)$ is the scattering operator for the portion of the slab from depth a to b.

We can simplify this further by replacing the $\mathrm{e}^{-c\Delta z}$ term with $1-c\Delta z+o(\Delta z^2)$ and taking advantage of the fact that for a layer that is infinitesimally thin [GY89],

$$\mathbf{S}(z,z+\Delta z) = \mathbf{k}\Delta z + o(\Delta z^2).$$

Making these simplifications and then taking the difference between the new scattering

operator $\mathbf{S}(z+\Delta z)$ and the scattering operator of the original layer $\mathbf{S}(z)$, we have

$$\mathbf{S}(z+\Delta z)-\mathbf{S}(z) = o(\Delta z^2)+\Bigg(-\sigma_t\left(\frac{1}{\mu_i}+\frac{1}{\mu_o}\right)\mathbf{S}(z)+ \\ \big(\mathbf{k}(z,z+\Delta z)+\mathbf{k}(z,z+\Delta z)\mathbf{S}(z)+\mathbf{S}(z)\mathbf{k}(z,z+\Delta z)+ \\ \mathbf{S}(z)\mathbf{k}(z,z+\Delta z)\mathbf{S}(z)\big)\Bigg)\Delta z$$

Now we divide by Δz and take the limit as $\Delta z \to 0$, which gives us the infinitesimal change in the scattering function due to the addition of the new layer.

$$\frac{\partial \mathbf{S}}{\partial z} = -\sigma_t\left(\frac{1}{\mu_i}+\frac{1}{\mu_o}\right)\mathbf{S}+\big(\mathbf{k}+\mathbf{kS}+\mathbf{Sk}+\mathbf{SkS}\big) \tag{3.3}$$

The first approach to solving this equation came from Ambarzumian, who observed that for a semi-infinite atmosphere (i.e. one with a depth of ∞), $\frac{\partial \mathbf{S}}{\partial z}$ must be zero. In other words, given a semi-infinite homogeneous atmosphere, any finite thickness can be removed from it will leave the atmosphere's scattering function unchanged.[1] Thus,

$$0 = -\sigma_t\left(\frac{1}{\mu_i}+\frac{1}{\mu_o}\right)\mathbf{S}+\big(\mathbf{k}+\mathbf{kS}+\mathbf{Sk}+\mathbf{SkS}\big)$$

so

$$\begin{aligned}\mathbf{S} &= \frac{1}{\sigma_t\left(\frac{1}{\mu_i}+\frac{1}{\mu_o}\right)}(\mathbf{k}+\mathbf{kS}+\mathbf{Sk}+\mathbf{SkS}) \\ &= \frac{1}{\sigma_t}\frac{\mu_i\mu_o}{\mu_i+\mu_o}(\mathbf{k}+\mathbf{kS}+\mathbf{Sk}+\mathbf{SkS})\end{aligned}$$

For the finite thickness setting, however, a boundary condition is needed in order to convert this non-linear integro-differential equation into an integral equation. If we assume

[1] Ambarzumian's original observation was for the one-dimensional case, with only isotropic scattering in the atmosphere. Others later extended this to the anisotropic case.

that the object is bounded by a perfect absorber from below — i.e. $\mathrm{S}(0) = 0$ — then application of the Laplace transform gives Equation 3.4. General boundary conditions are most easily handled with the adding equations that are introduced in the next chapter.

$$\mathbf{S}(z) = \int_0^z \mathrm{e}^{-\sigma_t(1/\mu_i+1/\mu_o)(z-z')} \big(\mathbf{k}(z') + \mathbf{k}(z')\mathbf{S}(z') + \mathbf{S}(z')\mathbf{k}(z') + \mathbf{S}(z')\mathbf{k}(z')\mathbf{S}(z')\big)\,\mathrm{d}z'. \quad (3.4)$$

Writing Equation 3.4 with the operators expanded out, we have

$$\begin{aligned}\mathrm{S}(z, r_i \to r_o) = \int_0^z \mathrm{e}^{-(\sigma_t(x_i)/\mu_i+\sigma_t(x_o)/\mu_o)(z-z')} \Bigg(& \mathrm{k}(r_i(z') \to r_o(z')) + \\ & \frac{1}{4\pi}\int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} \mathrm{k}(r_o \to -r')\,\mathrm{S}(z', r_i \to r')\,\frac{\mathrm{d}r'}{\mu_{r'}^2} + \\ & \frac{1}{4\pi}\int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} \mathrm{S}(z', r' \to r_o)\,\mathrm{k}(r_i \to r')\,\frac{\mathrm{d}r'}{\mu_{r'}^2} + \\ \frac{1}{16\pi^2}\int_{\mathcal{R}^+_{\mathcal{M}^2(z')}}\int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} & \mathrm{S}(z', r'' \to r_o)\,\mathrm{k}(-r' \to -r'')\,\mathrm{S}(z', r_i \to r')\,\frac{\mathrm{d}r'}{\mu_{r'}^2}\frac{\mathrm{d}r''}{\mu_{r''}^2}\Bigg)\mathrm{d}z' \end{aligned} \quad (3.5)$$

where the ray $r(t)$ is a new ray along the same line as r, constructed by offsetting the origin by distance t along the z axis and $x_i = x(r_i(z-z'))$ and $x_o = x(r_o(z-z'))$.

This is a formidable equation, but like the equation of transfer, it expresses a simple fact about light scattering. With computers and numerical methods, it can be solved. We will discuss previous solution techniques and some new Monte Carlo approaches for solving it in Section 3.3.

3.2.2 Transmission

A similar derivation based on adding an infinitesimally thin layer and finding the change in transmission through the region of space gives the integral scattering equation for transmission, where the two rays enter on opposite sizes with respect to the z parameterization. We will omit the details of the derivation here.

$$\begin{aligned}
\mathrm{S}(z, r_i \to r_o) = \delta(x(r_i) + (x(r_o) + z\omega(r_o)/\mu_{r_o})T_r(x(r_i), x(r_o)) \times \\
\int_0^z \mathrm{e}^{-(\sigma_t(x_i)/\mu_i + \sigma_t(x_o)/\mu_o)(z-z')} \Bigg(\mathrm{k}(r_i(z') \to r_o(z')) + \\
\frac{1}{4\pi} \int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} \mathrm{k}(r_o \to -r')\, \mathrm{S}(z', r_i \to r')\, \frac{\mathrm{d}r'}{\mu^2_{r'}} + \Bigg) \mathrm{d}z' + \\
\int_0^z \mathrm{e}^{-(\sigma_t(x_i)/\mu_i + \sigma_t(x_o)/\mu_o)(z-z')} \Bigg(\frac{1}{4\pi} \int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} \mathrm{S}(z', r' \to r_o)\, \mathrm{k}(r_i \to r')\, \frac{\mathrm{d}r'}{\mu^2_{r'}} + \\
\frac{1}{16\pi^2} \int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} \int_{\mathcal{R}^+_{\mathcal{M}^2(z')}} \mathrm{S}(z', r'' \to r_o)\, \mathrm{k}(-r' \to -r'')\, \mathrm{S}(z', r_i \to r')\, \frac{\mathrm{d}r'}{\mu^2_{r'}} \frac{\mathrm{d}r''}{\mu^2_{r''}} \Bigg) \mathrm{d}z' \quad (3.6)
\end{aligned}$$

3.2.3 One-dimensional setting

There are useful special cases of the general scattering equations in the one-dimensional setting; this was where they were first derived. In one dimension, position in x and y is irrelevant, so the delta functions in the integral from the phase function disappear, leading to simpler formulas and easier implementation. A finite slab then has four scattering functions (see Figure 3.3): given illumination at the top, one gives the amount of light reflected at the top R^+ and another gives the amount of light transmitted at the bottom T^- [Cha60]. The other two, R^- and T^+, give reflection and transmission due to light incident at the bottom. R^+ and T^- are given in Equations 3.7 and 3.8. Note that in general $\mathrm{R}^+ \neq \mathrm{R}^-$ and $\mathrm{T}^+ \neq \mathrm{T}^-$; this was first demonstrated by Preisendorfer [Pre58a] and Hovenier has investigated the relationship between R^+ and R^- and T^+ and T^- in the general case in the

presence of polarized light [Hov69].

$$\begin{aligned}\mathrm{R}^+(z,\omega_i\to\omega_o) = \int_0^z \mathrm{e}^{-\sigma_t(z')(1/\mu_i+1/\mu_o)(z-z')}\,\sigma_s(z')\Bigg(&\mathrm{p}(z',\omega_i\to\omega_o)+\\ &\frac{1}{4\pi}\int_\Omega \mathrm{p}(z',-\omega'\to\omega_o)\,\mathrm{R}^+(z',\omega_i\to\omega')\,\frac{\mathrm{d}\omega'}{\mu'}+\\ &\frac{1}{4\pi}\int_\Omega \mathrm{R}^+(z',\omega'\to\omega_o)\,\mathrm{p}(z',\omega_i\to-\omega')\,\frac{\mathrm{d}\omega'}{\mu'}+\\ &\frac{1}{16\pi^2}\int_\Omega\int_\Omega \mathrm{R}^+(z',\omega''\to\omega_o)\,\mathrm{p}(z',-\omega'\to-\omega'')\,\mathrm{R}^+(z',\omega_i\to\omega')\,\frac{\mathrm{d}\omega'\mathrm{d}\omega''}{\mu'\mu''}\Bigg)\mathrm{d}z' \quad (3.7)\end{aligned}$$

$$\begin{aligned}\mathrm{T}^-(z,\omega_i\to\omega_o) = \delta(\mu_i-\mu_o)\mathrm{e}^{-\int_0^z\sigma_t(z')\mathrm{d}z'/\mu_i}+&\\ \int_0^z \mathrm{e}^{-\sigma_t(z)(z-z')/\mu_i}\,\sigma_s(z')\Bigg(&\mathrm{e}^{-z'/\mu_o}\mathrm{p}(z',\omega_i\to-\omega_o)+\\ &\mathrm{e}^{-z'/\mu_o}\frac{1}{4\pi}\int_\Omega \mathrm{p}(z',-\omega'\to\omega_o)\,\mathrm{R}^+(z,\omega_i\to\omega')\frac{\mathrm{d}\omega'}{\mu'}+\\ &\frac{1}{4\pi}\int_\Omega \mathrm{T}^-(z',\omega'\to\omega_o)\,\mathrm{p}(z',\omega_i\to-\omega')\frac{\mathrm{d}\omega'}{\mu'}+\\ &\frac{1}{16\pi^2}\int_\Omega\int_\Omega \mathrm{T}^-(z',\omega''\to\omega_o)\,\mathrm{p}(z',-\omega'\to-\omega'')\,\mathrm{R}^+(z',\omega_i\to\omega')\frac{\mathrm{d}\omega'\mathrm{d}\omega''}{\mu'\mu''}\Bigg)\mathrm{d}z' \quad (3.8)\end{aligned}$$

Note that the T equation is written in terms of R; this can make its solution somewhat inconvenient. Hovenier has reduced this pair of functions to a single *exit function*; particular values of R and T can both be computed directly from it [Hov77].

The application of the reflection and transmission operators to a function f is

$$(\mathbf{R}f)(\omega'\to\omega) = \frac{1}{4\pi}\frac{1}{\mu}\int_\Omega \mathrm{R}(\omega'\to\omega)\,f(\omega')\,\mathrm{d}\omega'$$

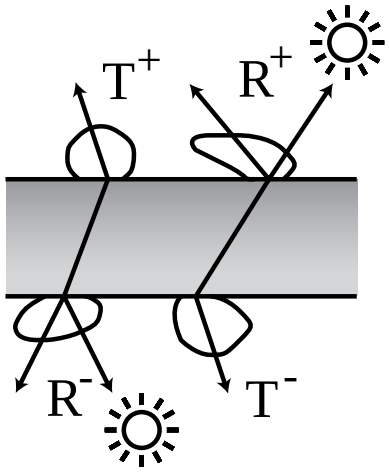


Figure 3.3: The two reflection and transmission functions of a slab.

which gives us a nearly familiar equation for computing reflected radiance at a point:

$$\mathrm{L_o}(\omega) = \frac{1}{4\pi}\frac{1}{\mu}\int_\Omega \mathrm{R}(\omega' \to \omega)\,\mathrm{L_i}(\omega')\,\mathrm{d}\omega'$$

The reflection function of a surface is thus related to its BRDF, f_r, by

$$\mathrm{R}(\omega' \to \omega) = 4\pi f_\mathrm{r}(\omega' \to \omega)\mu'\mu.$$

3.2.4 Discussion

The scattering equations thus bring us to a new approach for considering rendering problems. Note that there are no fundamentally new types of rendering problems that the scattering equations make accessible: as noted in Section 2.5.2, the formal inverse of the operator rendering equation can be used to solve the same kinds of scattering problems as well. For example, Hanrahan and Krueger effectively used a Neumman series expansion of the inverse to estimate four-dimensional scattering functions to compute BRDFs. This method can be extended to higher-dimensional scattering problems, and more sophisticated Monte Carlo techniques can be applied.

Conversely, as explained in Section 2.5, the scattering equation can be used for more than just pre-computing scattering functions. Given knowledge of particular viewing conditions, particular lighting conditions, or both, we can directly compute estimates of integrals such as $\mathbf{S}\mathrm{L}_e$ (where L_e is emitted radiance), rather than first computing $\mathbf{S}$ and then passing emitted light through it. Since both approaches are based on formulae that directly describe the physics of light scattering, it is not surprising that the two approaches are connected in this way. In fact (and reassuringly), the scattering equation can be derived directly from the equation of transfer [Pre65]. The advantage of the scattering function approach, however, is that it makes possible different flavors of solution methods that can be more efficient than approaches based on the equation of transfer.

3.3 Monte Carlo Solution

A previously uninvestigated technique for solving the scattering equations is Monte Carlo integration. Monte Carlo is a particularly effective technique for solving high dimensional integrals and integrals with discontinuities in the integrand. Its generality makes it possible to compute integrals where the functions in the integrand vary almost arbitrarily [KW86]; here, it allows wide variety in the possible phase functions, scattering and attenuation coefficients, and geometric shapes rendered.

In this section we first introduce basic principles of Monte Carlo integration. We then describe our solution technique for the one-dimensional reflection and transmission equations in order to motivate the general case. We explain the extension of this method to the three-dimensional case and discuss new difficulties that arise from the transition to this more general setting. General introductions to Monte Carlo integration include books by Spanier and Gelbard [SG69], Kalos and Whitlock [KW86] and Fishman [Fis96].

3.3.1 Monte Carlo Overview

We will briefly introduce notation and summarize basic principles of Monte Carlo integration. Monte Carlo is a statistical integration technique based on taking random samples from the domain of an integrand and using them to estimate the value of an integral. The standard Monte Carlo estimator gives the value of an integral $\int_D f(x)\mathrm{d}x$ over some domain D as the expected value of

$$\frac{1}{N}\sum_{i}^{N}\frac{f(X_i)}{p(X_i)},$$

where the X_i are random variables drawn from a sampling distribution $p(x)$ over D. This distribution must be non-zero for all x where $f(x) > 0$.

The most straightforward approach is to use a uniform distribution over D. However, given some prior knowledge about the shape of the integrand $f(x)$, a more efficient approach is to use *importance sampling*, where the distribution $p(x)$ is more likely to generate samples where $f(x)$'s value is relatively large. Importance sampling is one of a family of techniques for variance reduction that try to compute accurate estimates of the integral without just increasing the number of samples taken, N.

Another important technique for improving the efficiency of Monte Carlo integration is *Russian roulette*, which makes it possible to skip the computational expense of evaluating terms that are likely to make a small contribution to the overall result. With probability q, an estimate $f(X_i)/p(X_i)$ is not evaluated and is instead replaced with a value c. Otherwise, the estimate is still computed but is replaced with

$$\frac{f(X_i)/p(X_i) - qc}{1-q}.$$

The expected value of the estimate is still the value of the integral $\int_D f(x)\mathrm{d}x$. There is great flexibility in how the values q and c are chosen; if the function $f(x)$ is expensive to evaluate or can sometimes be closely approximated with a constant term, Russian roulette

can substantially improve the efficiency of Monte Carlo integration.

3.3.2 Solution in One Dimension

We will describe a simple recursive solution of the integral scattering equation. Because $\mathbf{S}(z)$ in Equation 3.4 is written recursively in terms of integrals of scattering functions of $\mathbf{S}(z')$, we can evaluate an estimate of $\mathbf{S}(z)$ based on a Monte Carlo integration process. The general form of the integral to estimate the value of is

$$\int_0^z e^{-f(z')} g(z') \mathrm{d}z',$$

where $f(z')$ and $g(z')$ represent corresponding terms of Equation 3.5. This process is most easily understood in the one-dimensional case (Equation 3.7).

The standard Monte Carlo estimator says that we can compute this integral by sampling z_i values from some distribution $p(z)$ and computing the estimate

$$\frac{1}{N} \sum_i^N \frac{e^{-f(z_i)} g(z_i)}{p(z_i)}.$$

Of the terms in this expression, $e^{-f(z_i)}$ and $p(z_i)$ can be computed immediately. $g(z_i)$ is itself a sum of four terms, one of which is easily computed (the phase function), and the rest of which involve single or double integrals. Monte Carlo integration can then be applied to estimate those integrals, as we will show shortly.

Sampling from exponential distributions

For the first sampling distribution of depths z_i, it is better to use a distribution that matches the $e^{-f(z)}$ term's distribution than to use a uniform distribution, particularly when the optical thickness of the scattering medium is large. For constant attenuation functions, the

exponential term's distribution can be sampled directly: to sample the integral

$$\int_0^z \mathrm{e}^{az'}\,\mathrm{d}z'$$

where $a = -\sigma_t(1/\mu_i + 1/\mu_o)$, we first find the probability density function

$$p(z') = \frac{a\,\mathrm{e}^{az'}}{\mathrm{e}^{az} - 1}.$$

The cumulative distribution function $P(z)$ is $\int_0^z \mathrm{p}(z')\mathrm{d}z'$. Given a random number ξ between 0 and 1, we set $\xi = P(z')$ and solve for z', giving

$$z' = \frac{\log(1 + \xi(\mathrm{e}^{az} - 1))}{a}.$$

More generally, if the attenuation term isn't constant but varies with depth, the a sampling distribution that matches it cannot in general be computed analytically. In this case, we can sample an optical thickness and march through the medium until that distance has been covered. In either case, the resulting sample is weighted by the exponential term at z' divided by the pdf.

3.3.3 Estimating the Hemispherical Integrals

Similar Monte Carlo integration machinery can be applied to compute the integrals in the $g(z)$ term. For example, consider the second term of the reflection function integral,

$$\frac{1}{4\pi}\int_\Omega \mathrm{p}(z', -\omega' \to \omega_o)\,\mathrm{R}^+(z', \omega_i \to \omega')\,\frac{\mathrm{d}\omega'}{\mu'}.$$

The values of z', ω_i, and ω_o are given; values for ω' are sampled from a sampling distribution to compute the estimate

$$\frac{1}{N}\sum_{i}^{N}\frac{\mathrm{p}(z',-\omega_j\rightarrow\omega_o)\mathrm{R}^+(z',\omega_i\rightarrow\omega_j)\frac{1}{\mu'}}{p(\omega_j)}.$$

As before, a uniform distribution of directions over the hemisphere could be used. It is not feasible to directly sample from the distribution of $\mathrm{p}(z',-\omega\rightarrow\omega_o)\mathrm{R}^+(z',\omega_i\rightarrow\omega)$, as doing so effectively requires being able to compute the integral analytically. However, sampling from the phase function's distribution or an approximation to R^+'s distribution can give better results than just using a uniform sampling distribution.

Sampling phase functions

Many phase functions can be sampled exactly; we will describe three in this section. Our general sampling routine takes an incident direction and two random numbers, ξ_1 and ξ_2, which are used to select an outgoing direction. The routine calls the appropriate lower-level sampling routine based on the phase function, constructs an outgoing vector, and returns the new direction as well as the probability of it being chosen.

The phase function-specific sampling routines take a single random number and return a value $\cos\theta$ that represents the cosine of the offset of the new vector from the original one (see Figure 3.4) as well as the probability density function for sampling that offset. The offset angle is enough to give us a circle of possible directions for the new vector. We then use the second random number to select a specific outgoing direction by choosing a point around the circle.

An isotropic phase function is most easily sampled. In order to sample with a uniform density over the sphere of possible directions, we compute $\cos\theta = 1-2\xi$. By uniformly sampling in $\cos\theta$ (and thus non-uniformly sampling in θ), this accounts for the $\sin\theta$ effect of the measure on the sphere $\sin\theta d\theta d\phi$ [CW93].

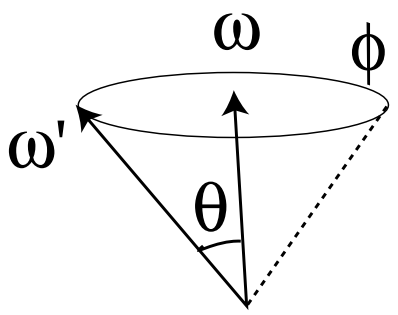


Figure 3.4: Given an incident vector ω, a scattered vector ω' is sampled via a two-step process. First the phase function samples an offset angle θ based on its distribution. Then, ω' is computed using a second random number to sample a direction from the circle of possible directions given by θ.

The Henyey-Greenstein phase function (Equation 2.3) is also easily sampled [HK93]. By applying the standard integration and inversion approach to find a sampling distribution that is propoertional to the function's distribution, we have

$$-\cos\theta = \frac{1+\frac{g^2-(1-g^2)}{(1-g+2g\xi)^2}}{2g},$$

and where the probability density is just the value of the phase function for the direction chosen.

When sampling a weighted sum of phase functions, e.g. Equation 2.4, a two-stage process is used. First, we randomly choose one of the terms to sample (for example, with probability proportional to their relative weights). We then use that term's importance sampling routine to determine $\cos\theta$. However, to compute the pdf, we apply multiple importance sampling [VG95] (the balance heuristic in particular), and return instead the weighted sum of pdfs from each of the terms.

A slight complication is that we often need to sample a direction that is constrained to be in a particular hemisphere. For example, when computing estimates of the **kS** term of Equation 3.4, if we want to sample a direction based on **k**, that direction must be in

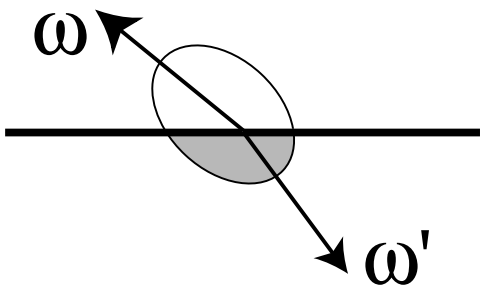


Figure 3.5: When sampling directions ω' for estimating the hemispherical integrals in the reflection equation, the directions ω' need to be in the opposite hemisphere (dark region) than the incident direction ω.

the downward hemisphere (see Figure 3.5). If the sampling method used sometimes does generate samples in the upward hemisphere, the value of the integrand must be taken as zero for such samples.

Sampling scattering functions

It is also useful to be able to use importance sampling to choose directions according to the scattering function. This is more difficult than sampling according to the phase function, since we don't know the value of the scattering function in closed form. Thus, it is necessary to come up with approximate importance functions that match the general shape of the scattering function well, but are still easily evaluated.

The simplest approach would be to treat the structure of the scattering function as completely unknown and use an isotropic sampling distribution. Because there is non-zero probability of sampling any direction, the value computed is guaranteed to be correct. However, a better approach tries to increase overall efficiency by sampling a distribution that better matches the scattering function's distribution.

We could do this by sampling according to a single-scattering approximation of the

scattering function. For optically thin objects and for objects with low albedos, most scattering is due to single scattering events, rather than chains of multiple scattering, anyway. Because we can sample phase functions exactly, this is slightly better, though it doesn't account for the multiple scattering encapsulated by the scattering functions.

We have developed a more robust sampling approach that accounts for what has been called *angular relaxation* [Yan97, Section 3.8]. It has widely been observed that after multiple scatterings, the original radiance distribution is eventually 'forgotten' and becomes completely isotropic. The larger the albedo, the greater this effect.

Yanovitskij has derived an expression that describes the isotropization due to multiple scattering events from the Henyey-Greenstein phase function [Yan97]. Expressed in our notation, the scattering distribution function after n scattering events is

$$\mathrm{p}(\omega \to \omega')^n = \frac{1-g^{2n}}{1+g^{2n}-2g|g^{n-1}|(-\omega\cdot\omega')^{3/2}}. \tag{3.9}$$

Using this improved distribution, an approximation to the scattering function's distribution is given by the infinite sum

$$\sum_{n=1}^{\infty} \alpha \mathrm{p}(\omega \to \omega') + \alpha^2 \mathrm{p}(\omega \to \omega')^2 + \cdots$$

Although this could be used for importance sampling by choosing a term to sample with probability proportional to its contribution, we can't apply the multiple importance sampling technique, as that would require that we compute the probability of *each* of the sampling strategies choosing the resulting direction. Therefore, we cut off the infinite sum after a small number of terms (e.g. three or four), and collect all remaining terms in a single perfectly isotropic importance function.

See Figure 3.6 for pseudo-code that summarizes the complete process for applying Monte Carlo integration to evaluate the 1D reflection function.

```
Procedure R(z, ω_i, ω_o) ≡
  (z′, pdf) := sampleDepth(σ_t, z)
  scale := e^{-(z-z')*σ_t*(1/μ_i + 1/μ_o)} / pdf
  result := σ_s * p(ω_i, ω_o)
  if (not terminate()) then
    (ω′, pdf') := sampleAngle(p, ω_o)
    result := result + σ_s * p(-ω′, ω_o) * R(z′, ω_i, ω′) /
          (cos ω′ * pdf')
    (ω″, pdf″) := sampleAngle(p, ω_i)
    result := result + R(z′, ω″, ω_o) * σ_s * p(ω_i, -ω″) /
          (cos ω′ * pdf″)
    result := result + R(z′, ω′, ω_o) * σ_s * p(-ω″, -ω′) *
          R(z′, ω_i, ω″) / (cos ω′ * cos ω″ * pdf' * pdf″)
  endif
  return result * scale
```

Figure 3.6: Pseudo-code for evaluation of the one-dimensional reflection equation. The phase function $p()$, σ_s, and σ_t are all potentially varying with depth. The terminate() function probabilistically stops the recursion using Russian roulette based on the weighted contribution that this estimate of R will make to the final solution. The sampleAngle() function uses importance sampling to choose an outgoing angle based on the phase function and the incoming angle; it returns the new direction and its probability density.

Discussion

There is an important difference between the process of estimating values of the integral scattering equation and Monte Carlo solutions of the equation of transfer: as the recursion continues, the z' at which we are estimating $\mathbf{S}$ is monotonically decreasing. Once we have chosen a depth at which to estimate $\mathbf{S}$, all scattering above z' is irrelevant; it has already been accounted for. In effect, we are able to make a single pass through the medium from top to bottom, peeling off layers and solving scattering problems for thinner sub-objects. In comparison, standard approaches to solving the equation of transfer such as those based on the Neumann series expansion do not create a progressively simpler problem as they proceed.

3.3.4 Solution in Three Dimensions

In the 3D case, this sampling process is less straightforward due to the delta function in the ray space phase function. Fortunately, delta functions generally fit easily into Monte Carlo sampling schemes. For example, given two rays r and r', the $\mathbf{k}(z')$ term of Equation 3.4 is zero unless both r and r' start at the same point *and* the z' depth sampled in the first sampling step above matches that point. In general two rays in 3D do not meet at all. Therefore, in the process of sampling the integrals, whenever we have a choice of rays to sample, sometimes we must carefully choose a ray and a depth such that this delta function is non-zero. To make this easier, we separate $\mathbf{S}$ into two components, $\mathbf{S}^{\mathrm{s}}$, scattering due to a single scattering event, and $\mathbf{S}^{\mathrm{m}}$, scattering due to multiple scattering events. This is analogous to distribution ray tracing with a mixed pure specular and diffuse surface where the two parts need to be sampled separately.

$$\mathbf{S}_{\mathrm{s}}(z) = \int_0^z \mathrm{e}^{\cdots}\mathbf{k}\,\mathrm{d}z'$$

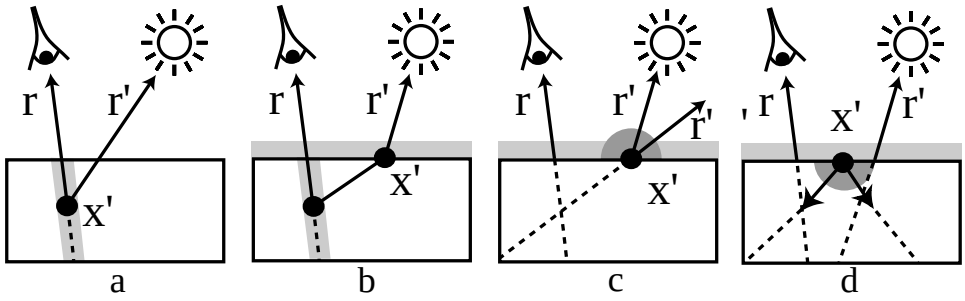


Figure 3.7: Sampling rays r' in the 3D case; hilighted regions denote terms that are free to be sampled. From left to right: for direct lighting, all scattering events are along r's path through the medium; for $\mathbf{S}^{\mathrm{s}}$, we sample a distance along r to find a scattering event—this gives a ray that connects through the point x' on the boundary; for $\mathbf{S}^{\mathrm{m}}$, we have more freedom to sample the direction of r'' and can use a variety of sampling distributions; finally, for **SkS**, we sample a point for the scattering event **k** and then sample the two outgoing directions.

and

$$\mathbf{S}_{\mathrm{m}}(z) = \int_0^z \mathrm{e}^{\cdots}(\mathbf{kS} + \mathbf{Sk} + \mathbf{SkS})\,\mathrm{d}z'.$$

Thus, $\mathbf{S} = \mathbf{S}^{\mathrm{s}} + \mathbf{S}^{\mathrm{m}}$.

Consider the specific case of estimating $\mathrm{L_o} = \mathbf{S}\mathrm{L}_e$ for a given outgoing ray r and a single point light source. Separating $\mathbf{S}$, we have two integrals, $\mathbf{S}^{\mathrm{s}}\mathrm{L}_e + \mathbf{S}^{\mathrm{m}}\mathrm{L}_e$. The first term is easily handled: it just represents single scattering of emitted light in the medium, so all scattering events are along r's path through the object. We choose positions for scattering events (i.e. x' in Figure 3.7a) using importance sampling to select points along r as above. Given these points, the incoming ray r' follows directly since the light is a point source; for an area light, a point can be chosen on the source and r then follows.

Moving on to $\mathbf{S}^{\mathrm{m}}\mathrm{L}_e$, we first consider the term

$$\mathbf{Sk}\mathrm{L}_e = \mathbf{S}^{\mathrm{s}}\mathbf{k}\mathrm{L}_e + \mathbf{S}^{\mathrm{m}}\mathbf{k}\mathrm{L}_e$$

(treatment of **kS** is analogous), we first randomly sample a point x' on the surface where

the **k** scattering event happens, using an exponential distribution centered around $x(r)$ (Figure 3.7b). This strategy is based on the assumption that the longer the distance light travels under the surface, the more it will be attenuated and the less impact it will have. This defines a ray r' to the light due to the point light assumption (as above, area lights are a straightforward extension). The second single scattering event must be along r's path through the medium and must have a direction such that it passes through $x(r')$ in order for all of the respective delta distributions to be non-zero. We therefore can again chose a depth along r with importance sampling.

There is more freedom in sampling from the $\mathbf{S}^{\mathrm{m}}\mathbf{k}\mathrm{L}_e$ term (Figure 3.7c). We choose a ray r' as above, and still must have the **k** scattering event at $x(r')$ for the delta distribution in **k** to be non-zero. However, the direction of r'' can be chosen arbitrarily since $\mathbf{S}^{\mathrm{m}}$ doesn't have the delta distribution along the path of r through the medium that $\mathbf{S}^{\mathrm{s}}$ does. We simply sample from the phase function's distribution based on $\omega(r')$ to get the ray direction for the r''.

Finally, the $\mathbf{SkS}\mathrm{L}_e$ term is slightly different: we also need to choose two rays that meet at a point where the **k** term will be evaluated (Figure 3.7d). We sample the shared ray origin from an exponential distribution centered around the midpoint between $x(r)$ and $x(r')$. Given this origin that the two new rays share, we again use importance sampling with the phase function to choose the two ray directions.

3.4 Results

In this section, we demonstrate the use of the 1D scattering equations to rendering complex surfaces and show applications of the 3D scattering equation to accurate rendering of surfaces, accounting for light that enters the surface some distance from where it exits.

3.4.1 Accuracy

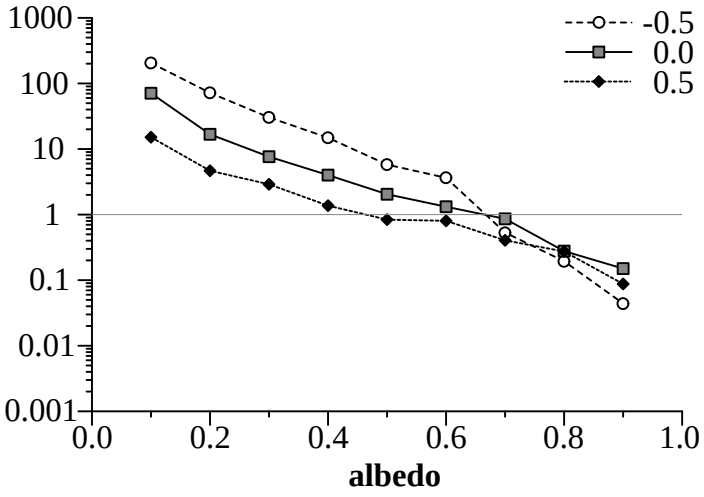
We tested our implementation's accuracy against a model that corresponds to what is known as "the standard problem" in astrophysics. The parameters to this model are an atmosphere's optical thickness, albedo, and phase function. The resulting scattering functions that describe the atmosphere's aggregate scattering function have been computed and tabularized by many authors. We compared our results to tables from Bellman et al. [BKP63], which have results computed by using Gaussian quadrature to generate a system of differential equations which were then solved via the Runge-Kutta method.

For a set of roughly forty randomly-selected albedos, thicknesses, and pairs of angles, we found excellent agreement with the scattering function values our routines computed. Finally, we verified that our 3D implementation gave the same results as the 1D equation for uniformly illuminated planar objects that have homogeneous scattering properties in xy.

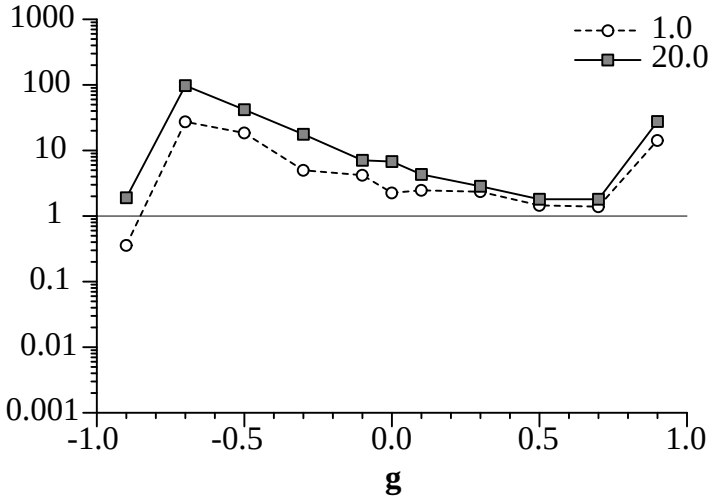
3.4.2 Efficiency

We then conducted a series of experiments to compare the efficiency of our solution method to a standard solution method that uses the equation of transfer. We implemented a Monte Carlo sampling routine that uses the equation of transfer to estimate the scattering function of a medium for a pair of angles based on a random walk. Our implementation is similar to the algorithm described by Hanrahan and Krueger [HK93]: a particle is injected into the medium from the incident direction and followed along a path through the medium. The walk is biased so that at each scattering event, the attenuation to the surface in the outgoing direction is computed and the result is accumulated to estimate the function's value. Russian roulette is used to terminate this process, based on the accumulated weight of the path.

After verifying that both methods converged to the same results, we compared their



Ratio of variance as a function of albedo
for various g. Depth = 2.



Ratio of variance as a function of g
for various depths. Albedo = 0.4

Figure 3.8: Comparing our solutions of the "standard problem." After giving each method the same amount of processor time to compute the best possible solution, we have graphed the ratio of variance when the equation of transfer is sampled to the variance when the scattering equation is sampled. Points above 1 on the *y* axis indicate situations where the scattering equation is more efficient. Because both sampling methods converge at the same rate asymptotically, the ratio of running time to compute solutions of equivalent quality is proportional to the variance ratio.

relative efficiency. For a variety of thicknesses, scattering coefficients, absorption coefficients, and phase functions, we computed accurate estimates of the scattering function for a pair of angles. We then applied both solution methods to computing estimates of the scattering functions for the pair of angles, giving each the same amount of processor time. The same Russian roulette termination parameters were used for each method and importance sampling was applied in analogous places (e.g. for sampling the outgoing direction of the phase function at scattering events for the equation of transfer). Our implementation generally computed five to ten estimates with the equation of transfer in the time it took to compute one estimate with the scattering equation.

The graphs in Figure 3.8 show some of the results. We computed the ratio of variance of the equation of transfer solution to the scattering equation solution, after giving each the same amount of processor time. For some cases scattering equation solution had 5 to 10 times less variance, and for some configurations (strongly anisotropic phase functions and very thick objects), it sometimes had over 100 times less variance. Although the scattering equation approach generally performed well, for cases with high albedos the equation of transfer was sometimes more efficient. The scattering equation also did well for most phase function parameters, except for extreme backward scattering. In this case, although most of incident light is quickly scattered back out of the top, we still continue to work through the z depth of the medium, not allocating effort as well as may be optimal.

3.4.3 Subsurface Scattering from Volumes

As a final example, we generated some images to demonstrate the use of the 3D scattering equation to compute reflection from complex surfaces and performed some experiments to understand the properties of subsurface light transport.

To determine how distance from the point of illumination affected the intensity of reflected light, we illuminated half of a slab from the direction along its normal and looked

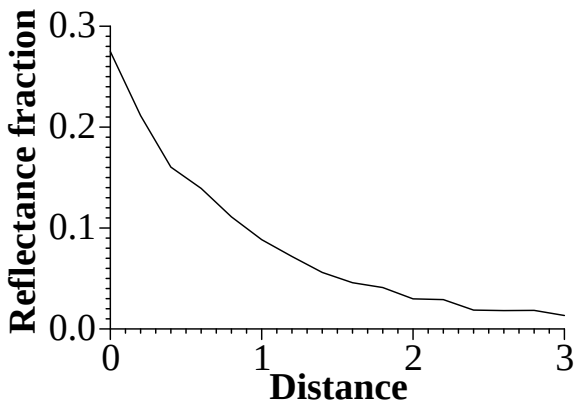


Figure 3.9: Reflection function magnitude in the unilluminated part of an object as a function of distance from the boundary of illuminated region. $\sigma_a = 0.5$, $\sigma_s = 0.5$, $g = 0$.

at the scattering function's magnitude in the normal direction at a series of points moving away from the illuminated area. Figure 3.9 shows the results. As one might expect, reflectance drops off roughly exponentially. Other experiments showed that as the object gets thinner, light entering from far away becomes less important, because more light scatters out of the object before traveling very far. These observations help validate some of the assumptions made in designing importance sampling techniques for 3D scattering equation.

We also rendered some images of marble, in the form of a marble block. Scattering properties were computed procedurally using noise functions [Per85]. Figure 3.10 shows a comparison of rendering a block with the 1D scattering equation (left) compared to a rendering with the 3D scattering equation (right). The right halves of the blocks were brightly illuminated by a directional light source, while the left halves were lit dimly. There are a number of significant differences between the two images. Most strikingly, when subsurface light transport is accounted for we can see the effect of light that entered in the illuminated half and then scattered into the unilluminated half. Furthermore, the veins of the marble, where the attenuation coefficient is high, cast shadows inside the volume; this effect is missing in the 1D case. A subtle difference between the two can be seen along

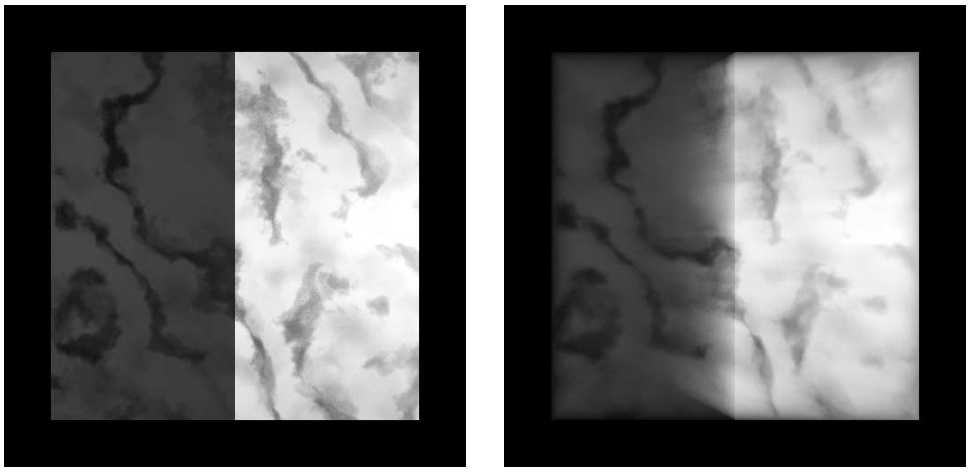


Figure 3.10: Comparison of rendering subsurface scattering from a side-lit marble cube with the 1D scattering equation (left) and the 3D scattering equation (right). Since the 3D solution considers light that enters the surface away from where it exits, subsurface light transport is more accurately modeled.

the edges: they are more transparent in the 3D version, since the geometry of the object is accounted for in computing subsurface scattering and rays leave the object after a short distance.

3.5 Previous Work

In the early 1900s, McCleland and McClelland and Hackett developed techniques similar to the invariant imbedding method used to derive the scattering equations here. After conducting a series of experiments to measure reflected flux from a variety of materials after bombardment from β rays, they derived a pair of coupled differential equations describing the change in upward and downward radiation at points in a one-dimensional medium [McC06, MH06a] (see also [McC05, MH06b]). These results were used to better estimate scattering and attenuation properties of objects, by distinguishing between singly and multiply scattered radiation.

Ambarzumian's description of the invariance of the reflection function of a semi-infinite

atmosphere was the next important innovation for this approach [Amb42, Amb43, Amb58]. He observed that the intensity and distribution of reflected light from a semi-infinite homogeneous atmosphere is unchanged as finite layers are added to or removed from it (see Section 3.2); this led to an entirely new integral scattering equation that described scattering from such an atmosphere directly, entirely avoiding the need to use the equation of transfer. Chandrasekhar refined this insight and placed it on a firm mathematical footing [Cha60]. He then developed generalized principles of invariance that led to integro-differential equations that described reflection and transmission from finite atmospheres.

Bellman and Kalaba were the first to derive the purely integral form of this equation for theoretical analysis of solutions to the scattering equation [BK56]. Further theoretical analysis of the equation in integral form as well as the generalization to depth-varying albedos was performed by Ueno [Uen60] and Busbridge [Bus61]. Redheffer [Red62] points out that similar ideas to Ambarzumian's have been applied for computing impedance in circuits. He also provides a wide set of references to similar work in other areas (for example, probability.)

3.5.1 Extensions to higher dimensions

There have been a number of previous efforts to generalize the scattering equations to higher dimensions. Bellman et al. derived a scattering function that accounted for variation in the x direction as well as in the z direction [BKU62] and Corones has derived a system of equations for scattering in a two-dimensional medium, under the assumption that scattering happened at discrete points and in discrete directions [Cor89]. Faber et al. investigated the generalization of the adding equations to two dimensions [FSW89], and Nelson et al. derived non-linear integro-differential equations describing reflection and transmission of radiation in a two-dimensional medium, neglecting angular distribution of radiation [NSV92]. More recently, Nelson and Seth have derived scattering equations in

the two-dimensional case where the particles are allowed to move in only four directions [NS97]. In some sense this is a generalization of the Kubelka-Munk model to two dimensions. Tessendorf has developed a method of invariant imbedding for general geometries [Tes90, Tes87].

Recently, Wang has derived a scattering equation in the three-dimensional case where incident illumination from a distant source is constant over the entire upper boundary of the region and where the phase function varies only in z [Wan90] [NUW98, Section 4.6]. Thus far, it has only been applied to the problem of estimating the reflectance of an object on the ground, given information about reflected light from an atmosphere above the object [Wan93, WU94].

3.5.2 Application to computer graphics

The first graphics researchers to recognize the importance of the scattering equations were Max et al. [MMKW97]. They used the one-dimensional scattering equation to compute light scattering in tree canopies by deriving a system of ordinary differential equations. They solved this system with an adaptive Runge-Kutta method, using a clever application of the Fourier transform to avoid an $O(n^3)$ matrix multiplication. Because their solution technique discretizes the hemisphere into sets of angles, it becomes increasingly expensive for strongly peaked phase functions. More importantly, the viability of the extension of this solution method to 3D scattering problems has not been demonstrated.

3.5.3 Previous Solution Methods

Techniques previously used to solve the scattering and adding equations have been based on the integro-differential form such as Equation 3.3. Two comprehensive reviews are given by Hansen and Travis [HT74] and Irvine [Irv75]. See also van de Hulst's book [van80] and Peraiah's survey article for recent developments in and comparison of various solution

methods [Per99].

Gaussian Quadrature

The most widely used technique thus far has been to apply Gaussian quadrature (see e.g. [Atk93] for a general introduction to numerical analysis and quadrature methods) to Equation 3.7, in integro-differential form:

$$\begin{aligned}\frac{\partial \mathrm{R}(z,\omega_i\rightarrow\omega_o)}{\partial z} = -\sigma_t(z)\left(\frac{1}{\mu_i}+\frac{1}{\mu_o}\right)\mathrm{R}(z,\omega_i\rightarrow\omega_o)+\mathrm{p}(z,\omega_i\rightarrow\omega_o)+\\ \frac{1}{4\pi}\int_\Omega \mathrm{p}(z,-\omega'\rightarrow\omega_o)\,\mathrm{R}^+(z,\omega_i\rightarrow\omega')\,\frac{\mathrm{d}\omega'}{\mu'}+\\ \frac{1}{4\pi}\int_\Omega \mathrm{R}^+(z,\omega'\rightarrow\omega_o)\,\mathrm{p}(z,\omega_i\rightarrow-\omega')\,\frac{\mathrm{d}\omega'}{\mu'}+\\ \frac{1}{16\pi^2}\int_\Omega\int_\Omega \mathrm{R}^+(z,\omega''\rightarrow\omega_o)\,\mathrm{p}(z,-\omega'\rightarrow-\omega'')\,\mathrm{R}^+(z,\omega_i\rightarrow\omega')\,\frac{\mathrm{d}\omega'\mathrm{d}\omega''}{\mu'\mu''}\Bigg)\end{aligned} \quad (3.10)$$

By applying the quadrature formula

$$\int_0^1 f(x)\mathrm{d}x = \sum_{i=1}^{N} w_i f(x_i)$$

to the integrals, where the weights w_i and evaluation points x_i come from the Gaussian quadrature rules, a system of non-linear ordinary differential equations can be constructed. If we write $\mathrm{R}(z,\omega_i\rightarrow\omega_j)$ as $f_{ij}(z)$, we have a set of N^2 ordinary differential equations:

$$\begin{aligned}\frac{\partial f_{ij}(z)}{\partial z} = -\sigma_t(z)\left(\frac{1}{\mu_i}+\frac{1}{\mu_j}\right)f_{ij}(z)+\mathrm{p}(z,\omega_i\rightarrow\omega_j)+\\ \frac{1}{4\pi}\sum_{k=1}^{N}\mathrm{p}(z,-\omega_k\rightarrow\omega_j)\,\frac{w_k}{\mu_k}f_{ik}(z)+\frac{1}{4\pi}\sum_{k=1}^{N}\frac{w_k}{\mu_k}f_{kj}(z)\,\mathrm{p}(z,\omega_i\rightarrow-\omega_k)+\\ \frac{1}{16\pi^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\frac{w_k w_l}{\mu_k\mu_l}f_{lj}(z)\mathrm{p}(z,-\omega_k\rightarrow-\omega_l)\,f_{ik}(z)\end{aligned} \quad (3.11)$$

Given the set of initial conditions $f_{ij}(0) = 0$, this nonlinear initial value problem can be solved with standard techniques, such as Runge-Kutta integration.[2] The computational expense of this process can be reduced by making use of the symmetry that $f_{ij}(x) = f_{ji}(x)$, though the asymptotic complexity is still $O(n^2)$. (Max et al.'s solution technique was based on these techniques, using a refinement suggested by Hansen and Travis [HT74, Section 3] to avoid an $O(n^3)$ matrix multiplication in the solution).

The advantage of this method is that it easily handles inhomogeneity in the medium, both with depth-varying albedos and depth-varying phase functions. The implementation is also relatively straightforward and can be based on libraries of numerical routines. However, the method breaks down in the face of complexity in the scattering medium: if the phase function is highly anisotropic, the quadrature rule must be applied with a large number of points N; with the $O(n^2)$ asymptotic growth of the method, it can quickly become inefficient. More significantly, the generalization of this technique to higher-dimensional settings quickly becomes intractable; we have not seen it applied to any problems more complex than computing values of the one-dimensional reflection and transmission equations.

The Doubling Method

The adding equations have been applied to computing scattering functions in a technique called the *doubling method*. Starting with a representation for the reflection and transmission functions of a very thin layer (this starting scattering function is often computed with a single-scattering approximation or with a successive orders method), the adding equations are applied to compute reflection and transmission functions for layers of progressively twice the thickness [Han71, van80]. For a homogeneous atmosphere, the adding equations

[2] This family of solution methods is often referred to in the literature as "invariant imbedding", thus confusing the method of deriving the equation to be solved with the numeric technique applied to solve it. In particular, many authors have failed to distinguish between the general and continuous scattering equations that can be derived using invariant imbedding and the approximate discrete forms of them that result from applying a quadrature-based solution method.

only need to be applied a logarithmic number of times in the overall thickness of the atmosphere; it can thus compute scattering functions of very thick atmospheres relatively quickly.

Some method must be used to compute the integrals for the adding equations; typically the reflection and transmission functions are projected onto some basis such that the adding equations can be evaluated directly via matrix inversion (e.g. [PP51]). In addition to efficiency for thick atmospheres, another advantage of the doubling method is that it doesn't directly depend on the scattering properties of the atmosphere (e.g. that it be isotropic). However, there are similar issues as with quadrature methods where techniques based on a discretization of directions require progressively finer discretizations as phase functions become more anisotropic.

Finally, the doubling method works poorly when the atmosphere is inhomogeneous. In that case it degenerates to the *adding method*, which takes a linear number of adding computations, rather than a logarithmic number.

Asymptotic fitting

Many methods that are usually efficient have difficulty with slabs with large optical thicknesses. Yet thick homogeneous slabs have reflection functions that change slowly and smoothly with depth. This observation led to the *asymptotic fitting* method. The reflection function is computed for a series of thicknesses—e.g. $R(x)$, $R(2x)$, and $R(4x)$—and is stored with some discrete representation. A curve is then fit to the scattering functions, and new scattering functions for thicker slabs can be extrapolated at very large thicknesses. This method has been applied by van De Hulst [van71] in conjunction with doubling method. See also van De Hulst's book for further discussion [van80] and Yanovitskij for more recent applications [Yan97].

Monte Carlo

Monte Carlo solution of the equation of transfer has been done by many authors; we are not aware of previous application to the scattering equations. An early example is Plass and Kattawar's investigation of reflection and transmission of polarized light from cloud models [PK68, KP68]. Somewhat more recently, Meier and Lee have investigated scattering from plane parallel atmospheres where energy absorbed at some wavelengths is emitted at others (i.e. fluorescence) [ML78, LM80, ML81].

Chapter 4

The Adding Equations

The adding equations make it possible to express scattering functions of aggregate objects or regions of space in terms of the scattering functions of the objects inside the region. Given two objects with known scattering functions, we often want to compute the new scattering function that describes overall scattering from the two objects together. This new scattering function must account for scattering from each of the objects as well as all scattering exchanges between the two. The adding equations give the appropriate theoretical basis for computing these new scattering functions correctly. They were developed in the fields of neutron transport and radiative transfer to solve just this problem [PP51, van80, TJH66] and were first introduced to computer graphics by Pharr and Hanrahan [PH00].

Computing scattering functions with the adding equations can often be done much more efficiently than by recomputing the scattering functions of the aggregate object from scratch [van80]. This stems from the fact that the scattering functions of the two objects, $\mathbf{S}_a$ and $\mathbf{S}_b$, already incorporate all of the multiple scattering events inside a and b, so it is only necessary to compute the effect of multiple scattering *between* the two objects. This can lead to a more efficient approach than solving for the overall scattering directly. After a few terms, the series usually converges quickly, as long as not too much of the light is

re-scattered at each step. Since the results of this computation are new scattering functions, they can themselves be used in further computations of new scattering functions.

In this chapter we will give a rigorous derivation of the adding equations in the rayspace setting based on the interaction principle. We will then describe new Monte Carlo methods for solving the adding equations, and we will demonstrate their use for rendering subsurface scattering from layered media.

4.1 The Interaction Principle

In the 1960s and 1970s, R. W. Preisendorfer investigated the relationship between the invariance principles that formed the basis of the scattering equations of the previous chapter and the classical equation of transfer [Pre65, Pre76]. He first demonstrated that the two world-views were equivalent; each could be derived from the other, and thus neither was more general or fundamentally more powerful than the other. Preisendorfer went on to state a more general *interaction principle* that subsumed both approaches to radiative transfer. He demonstrated that both could be derived directly from the interaction principle and showed a number of new approaches that it made possible.

4.1.1 Radiometric Invariance Principles

Classic work by George Stokes laid the foundation for the invariance and interaction principles of the Twentieth Century. In a classic nineteenth Century paper, he derived expressions for the amount of light reflected and transmitted from a stack of glass plates [Sto62]. Given reflection and transmission coefficients for an individual plate that describe the reflected flux at the boundary and directly transmitted flux through a homogeneous glass plate, ρ and τ, Stokes investigated the summation of all reflected and transmitted terms; he found

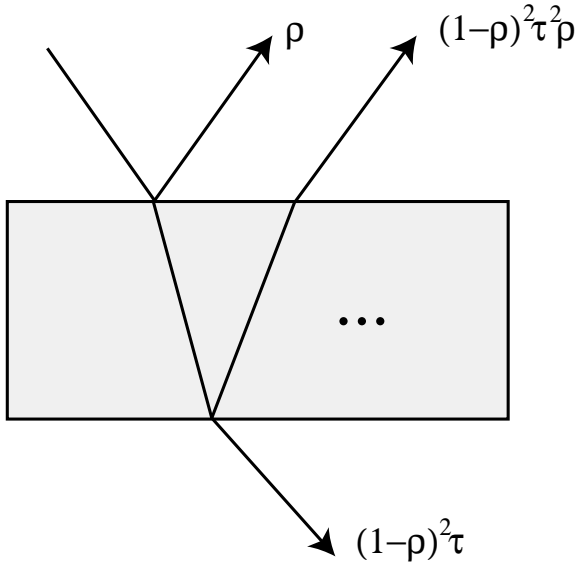


Figure 4.1: The overall reflection and transmission from a glass plate, expressed in terms of the reflection and transmission factors for single scattering or transmission events (after [Per99]).

the overall reflectivity and transmissivity to be

$$R = \rho + \tau^2(1-\rho)^2\rho(1+\rho^2\tau^2+\rho^4\tau^4+\cdots) = \rho\left(1+\frac{\tau^2(1-\rho)^2}{1-\rho^2\tau^2}\right)$$

and

$$T = (1-\rho)^2\tau(1+\rho^2\tau^2+\rho^4\tau^4+\cdots) = \frac{(1-\rho)^2\tau}{1-\rho^2\tau^2}$$

(See Figure 4.1).

Following similar lines, Stokes also showed that if a system of n and m plates is formed, where reflection and transmission values are known as R_n and T_n, etc, then the composite reflectance and transmittance are

$$R_{n+m} = R_n + R_m T_n^2(1+R_nR_m+R_n^2R_m^2+\cdots) = R_n + \frac{R_m T_n^2}{1-R_nR_m}$$

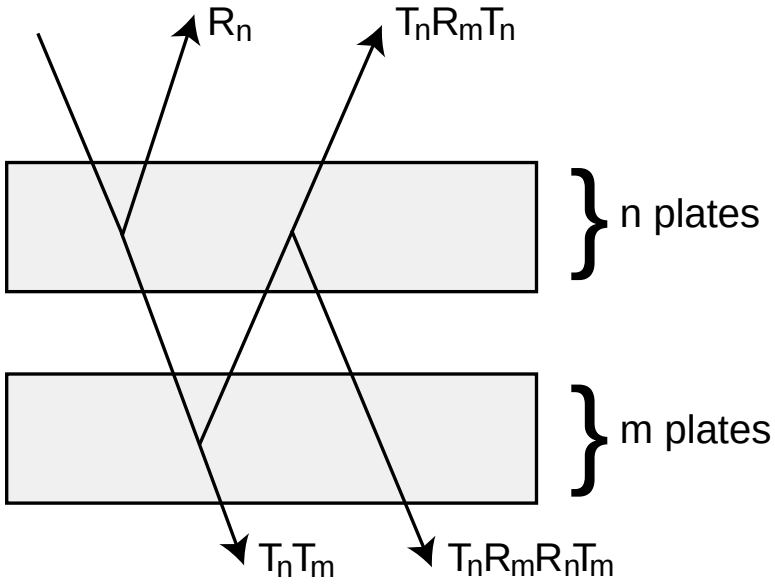


Figure 4.2: The overall reflection and transmission for a collection of n and m plates, expressed in terms of the reflection and transmission of each individually

and

$$T_{n+m} = T_n T_m (1 + R_n R_m + R_n^2 R_m^2 + \cdots) = \frac{T_n T_m}{1 - R_n R_m}$$

(See Figure 4.2). He thus introduced the observations that overall scattering could be computed directly in terms of the reflection and transmission functions of the individual layers, and that the reflection and transmission for two layers together could be computed based on the already-computed reflection and transmission of each one. The adding equations are effectively the generalization of these equations to encompass general operators and functions.

4.1.2 Statement and Use of the Interaction Principle

The interaction principle generalizes this approach to light scattering. It can be directly derived either from Preisendorfer's axiometric formulation of radiative transfer theory or directly from the electromagnetic theory [Pre65, Chapter XIV] [Pre76, Volume II, p. 189].

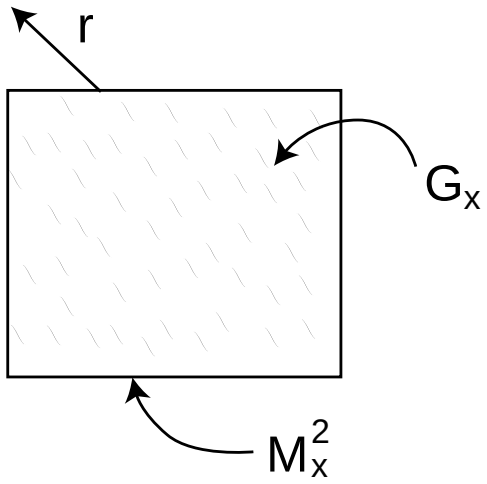


Figure 4.3: Basic ray space definition figure.

The basic property of radiative transfer theory that it relies on is the linearity of radiometric interactions.

As in the previous chapter, we use an instance of the ray space notation to describe our integration domains. Given some region of space G_x with a two-dimensional boundary given by a manifold $\mathcal{M}_x^2$, we define the set of all rays originating on $\mathcal{M}_x^2$ with direction leaving G_x by $\mathcal{R}_{\mathcal{M}_x^2} = \mathcal{M}_x^2 \times \Omega$. (see Figure 4.3 for an example of ray space.) We also define the differential measure of rays as in Chapter 3.

The scattering function S_x for a particular region G_x is defined over all rays in three-space; however it has a value of zero for all rays that do not originate on $\mathcal{M}_x^2$. This notational convention allows us to freely write products of scattering operators $\mathbf{S}_a\mathbf{S}_b$ where the boundaries $\mathcal{M}_a^2$ and $\mathcal{M}_b^2$ are not the same (giving a result that is identically the zero operator). See Section 3.1 for more on our ray space conventions. Incident and exitant radiance functions (e.g. L_i) are naturally defined over all rays in three-space, though we will only make use of their values on boundary manifolds $\mathcal{M}^2$.

4.1.3 Statement of the interaction principle

Quoting directly (with operator notation changed to match ours) the interaction principle states [Pre76, Section 3.2]:

> *For every* X, G, *A, B, m and n, for* X *is an optical medium and* G *is a subset of* X, *and* $A = (A_1, \ldots, A_m)$ *is a class of sets* A_i *consisting of incident radiometric functions on* G *and* $B = (B_1, \ldots, B_n)$ *is a class of sets* B_j *consisting of response radiometric functions on* G, *and m and n are positive integers, then there exists a unique set* $\{\mathbf{S}_{i\to j} : i = 1, \ldots, m, j = 1, \ldots, n\}$ *of linear (interaction) operators* $\mathbf{S}_{i\to j}$ *with domain* A_i *and range* B_j *with the property that for every element* $(a_1, \ldots, a_m)$ *of A there exists an element* $(b_1, \ldots, b_n)$ *of B such that*
>
> $$b_j = \sum_{i=1}^{m} \mathbf{S}_{i\to j}\, a_i \tag{4.1}$$
>
> *Or, in matrix form,*
>
> $$b = \mathbf{S}\, a \tag{4.2}$$

The most important element of the principle is that it asserts both the existence and uniqueness of the linear operators $\mathbf{S}_{i\to j}$. Because it is stated in terms of functions and operators, the interaction principle makes an exact statement about the system as formulated.

In most applications, the optical medium X is three-space, where the subsets G are convex surfaces[1] in X and A_i and B_j are radiance functions. Other common settings are the case where A_i and B_j are irradiance functions as well as the one-dimensional case, where interactions between plane-parallel media that are homogeneous in two dimensions are considered. When interaction between volumetric objects is being considered, two-dimensional surfaces are still generally used for the domains of A and B and the interactions are written in terms of incident and exitant radiance distributions over their boundaries.

[1] The principle holds for concave configurations as well, but the notation is more complex.

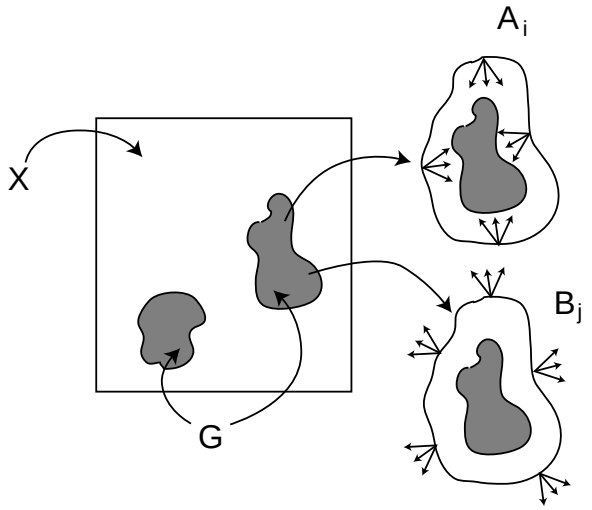


Figure 4.4: Basic setting for the interaction principle. Each subset G_i of the region has an incident radiance function A_i and an exitant radiance function B_i.

For the remainder of this chapter, we will consider X to be three-space and G to be two-dimensional boundaries of parallelipid-shaped regions of space. A_i is the set of incident radiance functions on these boundaries, and a_i is a particular radiance function $\mathrm{L_i}(x, \omega)$, where x is a point on the boundary and ω is an outgoing direction. B_j is the set of exitant radiance functions (see Figure 4.4).

4.1.4 Application and example

Application of the interaction principle generally involves a three-step process:

- First, the radiometrically interacting objects are enumerated: the subset G of X is identified and the incident radiometric quantities of interest are chosen, giving A_i. The outgoing radiometric quantities B_j then follow.

- Given the sets A_i and B_j, the interaction operators $\mathbf{S}_{i \to j}$ are determined. This lets us write the interaction matrix equation 4.2.

- We identify the radiometric relationships between the objects and the functions A_i

and B_j. In particular, by introducing additional equations that take into account radiometric principles (such as the invariance of radiance along lines in a vacuum), the system can be simplified.

The result of this is the system of operator equations to be solved. A variety of methods can then be applied, including projection onto a finite basis or Monte Carlo techniques. We only consider Monte Carlo approaches here.

4.2 Example: Interacting Plane Surfaces

We start with a simple example to show the use of the interaction principle, after [Pre76, Section 3.4]. Consider two planar surfaces X_1 and X_2, in three-space, with surface normals defined pointing toward the other surface (Figure 4.5). Each surface has an incident radiance distribution due to illumination from external sources (and not counting illumination from the other surface), given by L_i^1, L_i^2, respectively. Each surface has a scattering operator associated with it that describes how incident radiance is scattered by the surface, $\mathbf{S}_1$ and $\mathbf{S}_2$. The two surfaces are separated by a vacuum.

We wish to compute the outgoing radiance distribution accounting for the interaction of multiply-reflected radiance between the two of them. We apply the interaction principle, first defining the incident quantities:

- A_1: the incident radiance function on X_1 due to direct illumination, L_i^1.
- A_2: the incident radiance function on X_1 due to scattered light from X_2, $\mathrm{L}_i^{1\leftarrow 2}$.
- A_3: incident radiance on X_2 due to direct illumination, L_i^2
- A_4: incident radiance on X_2 due to light from X_1, $\mathrm{L}_i^{2\leftarrow 1}$.

Now we define the exitant quantities that we want to compute:

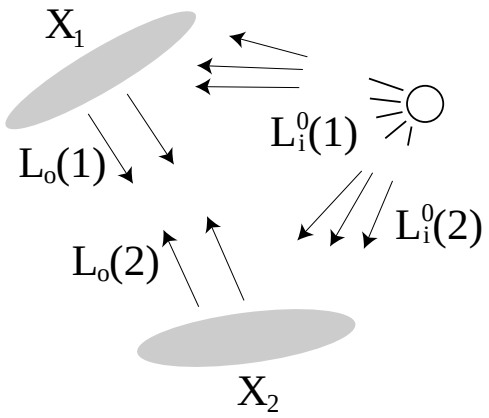


Figure 4.5: Given two surfaces X_1 and X_2, with incident radiance distribution functions L_i^1 and L_i^2, we apply the interaction principle to compute the outgoing radiance distributions L_o^1 and L_o^2.

- B_1: the exitant radiance distribution function on X_1, L_o^1.
- B_2: the exitant radiance distribution function on X_2, L_o^2.

Finally, we need to describe the interaction operators $\mathbf{S}_{ij}$. Rather than defining the four separate operators $\mathbf{S}_{11}$, $\mathbf{S}_{21}$, $\mathbf{S}_{32}$, and $\mathbf{S}_{42}$, we will just use general scattering operators $\mathbf{S}_1$ for X_1 and $\mathbf{S}_2$ for X_2. These scattering operators are defined for all points and all pairs of incident and exitant directions over each of the surfaces, and describe the amount of incident illumination scattered at each point from one direction to another.

From the interaction principle, we have

$$\begin{aligned} \mathrm{L}_o^1 &= \mathbf{S}_1 \mathrm{L}_i^1 + \mathbf{S}_1 \mathrm{L}_i^{1\leftarrow 2} \\ \mathrm{L}_o^2 &= \mathbf{S}_2 \mathrm{L}_i^2 + \mathbf{S}_2 \mathrm{L}_i^{2\leftarrow 1} \end{aligned}$$

Because the two objects are separated by a vacuum and because radiance is constant along lines in a vacuum, the incoming radiance at a point on X_1 from some direction due to outgoing radiance from X_2 is equal to the outgoing radiance at a corresponding point in a

corresponding direction on X_2. Thus,

$$\begin{aligned} \mathrm{L}_i^{1\leftarrow 2} &= \mathrm{L}_o^2 \\ \mathrm{L}_i^{2\leftarrow 1} &= \mathrm{L}_o^1 \end{aligned}$$

Therefore, we have

$$\begin{aligned} \mathrm{L}_o^1 &= \mathbf{S}_1\mathrm{L}_i^1 + \mathbf{S}_1(\mathbf{S}_2\mathrm{L}_i^2 + \mathbf{S}_2\mathrm{L}_i^{2\leftarrow 1}) \\ \mathrm{L}_o^2 &= \mathbf{S}_2\mathrm{L}_i^2 + \mathbf{S}_2(\mathbf{S}_1\mathrm{L}_i^1 + \mathbf{S}_1\mathrm{L}_i^{1\leftarrow 2}) \end{aligned}$$

If we eliminate L_o^1 and L_o^2 from the right side of the expressions, we then have

$$\mathrm{L}_o^1 = (\mathbf{S}_1\mathrm{L}_i^1 + \mathbf{S}_1\mathbf{S}_2\mathrm{L}_i^2)(\mathbf{I} - \mathbf{S}_1\mathbf{S}_2)^{-1} \tag{4.3}$$

$$\mathrm{L}_o^2 = (\mathbf{S}_2\mathrm{L}_i^2 + \mathbf{S}_2\mathbf{S}_1\mathrm{L}_i^1)(\mathbf{I} - \mathbf{S}_2\mathbf{S}_1)^{-1} \tag{4.4}$$

By expanding the inverted operator in a Neumann series, we come to the familiar result for Equation 4.3

$$\mathrm{L}_o^1 = \mathbf{S}_1\mathrm{L}_i^1 + \mathbf{S}_1\mathbf{S}_2\mathrm{L}_i^2 + \mathbf{S}_1\mathbf{S}_2\mathbf{S}_1\mathrm{L}_i^1 + \cdots .$$

In other words, the outgoing radiance from X_1 can be computed by computing outgoing radiance due to light scattered from direct illumination; then radiance due to scattered light from direct illumination on X_2 that is scattered towards X_1; direct illumination on X_1 that is scattered towards X_2, reflected back to X_1 and then scattered; and all further interactions of multiple scattering between the two objects.

This result shouldn't be surprising to anyone familiar with rendering-equation based approaches to solving light transport problems in computer graphics; what is noteworthy is the methodology that led us to this result, which completely sidesteps of the equation of transfer but has still led us to a rigorous result in the radiometry setting.

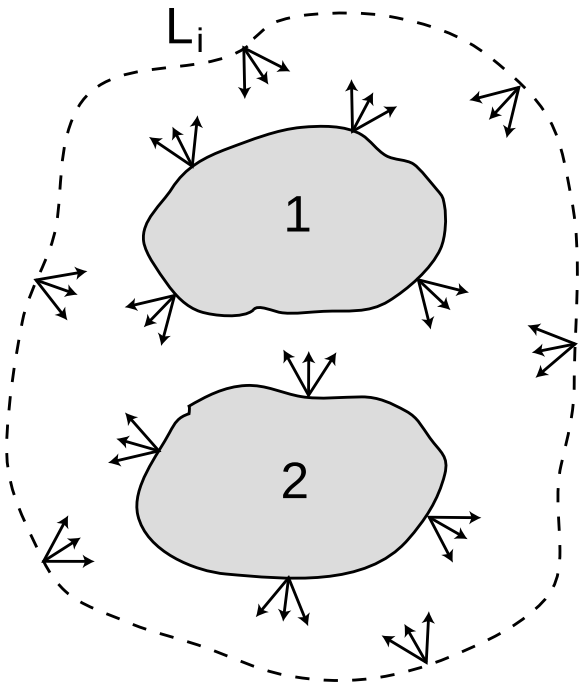


Figure 4.6: Basic setting deriving the adding equations using the interaction principle: two regions with known scattering properties are illuminated by an incident radiance function L_i. We'd like to find the overall scattering operator $\mathbf{S}_*$ that gives exitant radiance, $L_o = \mathbf{S}_* L_i$.

4.3 Derivation of Adding Equations

The original derivation of the adding equations by Peebles and Plesset followed a similar approach to the one used in the previous section, though the general interaction principle was not known at this time [PP51].

Consider two objects or regions of space in a volume, labeled 1 and 2 (see Figure 4.6). Applying the interaction principle in a similar manner, we can find that

$$\begin{aligned} L_o^1 &= \mathbf{S}_1 L_i + \mathbf{S}_1 L_i^{1\leftarrow 2} \\ L_o^2 &= \mathbf{S}_2 L_i + \mathbf{S}_2 L_i^{2\leftarrow 1} \end{aligned}$$

If the two objects abut or are separated by a vacuum, we can again describe incident radiance at one of them due to exitant radiance from the other.

$$\mathrm{L}_1 = \mathbf{S}_1\mathrm{L}_i + \mathbf{S}_1\mathrm{L}_2 \tag{4.5}$$

$$\mathrm{L}_2 = \mathbf{S}_2\mathrm{L}_i + \mathbf{S}_2\mathrm{L}_1 \tag{4.6}$$

Now we can substitute the second equation into the first, giving us

$$\mathrm{L}_1 = \mathbf{S}_1\mathrm{L}_i + \mathbf{S}_1(\mathbf{S}_2\mathrm{L}_i + \mathbf{S}_2\mathrm{L}_1)$$

Collecting L_1 terms, we have

$$(\mathbf{I} - \mathbf{S}_1\mathbf{S}_2)\mathrm{L}_1 = (\mathbf{S}_1 + \mathbf{S}_1\mathbf{S}_2)\mathrm{L}_i$$

Now we can compute the scattering function $\mathbf{S}_{12}$ that gives us outgoing radiance from the boundary of G_1 due to incident radiance L_i and scattering between G_1 and G_2: $\mathrm{L}_1 = \mathbf{S}_{12}\mathrm{L}_i$. Solving for $\mathrm{L}_1(\mathrm{L}_i)^{-1}$, we have

$$\mathrm{L}_1(\mathrm{L}_i)^{-1} = (\mathbf{I} - \mathbf{S}_1\mathbf{S}_2)^{-1}(\mathbf{S}_1 + \mathbf{S}_1\mathbf{S}_2) \tag{4.7}$$

$$= \mathbf{S}_1 + \mathbf{S}_1\mathbf{S}_2 + \mathbf{S}_1\mathbf{S}_2\mathbf{S}_1 + \cdots \tag{4.8}$$

$$= \mathbf{S}_1 + \mathbf{S}_1\mathbf{S}_2\mathbf{S}_{12} \tag{4.9}$$

A similar scattering function that gives the outgoing radiance function along the boundary of G_2 can be derived similarly. The sum of these two, i.e. the overall scattering function for the two objects that describes the overall outgoing radiance due to L_i, is

$$\mathbf{S}_* = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_1\mathbf{S}_2 + \mathbf{S}_2\mathbf{S}_1 + \mathbf{S}_1\mathbf{S}_2\mathbf{S}_1 + \cdots$$

Depending on the incident radiance distribution and the outgoing quantities of interest, many of these terms may be zero. For example, for the one-dimensional case of scattering layers where there is only incident illumination from the top and where we are only interested in outgoing radiance at the top, all terms ending with $\mathbf{S}_2$ disappear because L_i is zero over the boundary of the bottom region, thus $\mathbf{S}_2\mathrm{L}_i$ is zero. Furthermore, terms starting with $\mathbf{S}_2$ are also irrelevant, since the outgoing ray of interest starts on the top of the top layer and thus isn't on the bottom layer's boundary (recall the convention of Section 3.1).

4.3.1 One-dimensional setting

The adding equations are similarly simplified to integrals over just directions. In operator form, the scattering functions of two combined slabs *a* and *b* are

$$\begin{aligned}
\mathbf{R}^+_{a+b} &= \mathbf{R}^+_a + \mathbf{T}^+_a\mathbf{R}^+_b\mathbf{T}^-_a + \mathbf{T}^+_a\mathbf{R}^+_b\mathbf{R}^-_a\mathbf{R}^+_b\mathbf{T}^-_a + \cdots \\
\mathbf{R}^-_{a+b} &= \mathbf{R}^-_b + \mathbf{T}^-_b\mathbf{R}^-_a\mathbf{T}^+_b + \mathbf{T}^-_b\mathbf{R}^-_a\mathbf{R}^+_b\mathbf{R}^-_a\mathbf{T}^+_b + \cdots \\
\mathbf{T}^-_{a+b} &= \mathbf{T}^-_b\mathbf{T}^-_a + \mathbf{T}^-_b\mathbf{R}^-_a\mathbf{R}^+_b\mathbf{T}^-_a + \cdots \\
\mathbf{T}^+_{a+b} &= \mathbf{T}^+_a\mathbf{T}^+_b + \mathbf{T}^+_a\mathbf{R}^+_b\mathbf{R}^-_a\mathbf{T}^+_b + \cdots
\end{aligned}$$

The adding equations were developed by van De Hulst [van80], Twomey et al. [TJH66], and Redheffer [Red62] in the 1960s, and were applied to computing numerical results for isotropic scattering from plane-parallel atmospheres and scattering from clouds, respectively. They were first discovered in the field of neutron transport by Peebles and Plesset [PP51] and have since been applied to a wide variety of scattering problems. Later, Hansen extended the method to account for polarization [Han71]. The adding equations were later generalized to more general geometries by Preisendorfer [Pre76].

If one starts with a very thin slab with known reflection and transmission functions (if the slab is thin enough, a single scattering approximation works well), and then proceeds to

compute the scattering functions of slabs that are progressively twice as thick, the overall scattering functions can be computed entirely by this doubling method [Han71, van80]. For a homogeneous atmosphere, the adding equations only need to be applied a logarithmic number of times in the overall thickness of the atmosphere; it can thus compute scattering functions of very thick atmospheres relatively quickly. In addition to efficiency for thick atmospheres, another advantage of the doubling method is that it doesn't directly depend on the scattering properties of the atmosphere (e.g. that it be isotropic).

The Kubelka-Munk model for diffuse reflection from layered surfaces [KM31] uses a similar solution; if Equation 4.2 is written to operate on irradiance and if the scattering functions are isotropic, the Kubelka-Munk model can be easily derived. This is an instance of a *two-stream model*, similar to Schuster's [Sch05], which just considers the total amount of light going in the up and down directions, neglecting its distribution. The Kubelka-Munk model was first introduced to graphics by Haase and Meyer [HM92] and has been widely used. However, due to assumptions built into the model, either glossy specular reflection has to be ignored or multiple reflection between the specular component and the added layer is lost. A different approach to layer composition is due to Hanrahan and Krueger [HK93]; they compose scattering layers considering only one level of inter-reflection. This misses the effect of multiple internal reflections before light leaves the layer, which is important except for objects with very low albedos.

4.4 Monte Carlo Solution

Previous methods used to solve the adding equations have been based on projecting the scattering functions $\mathbf{S}_1$ and $\mathbf{S}_2$ onto a finite basis. This results in matrices that represent the operators, so that Equation 4.3 can be solved directly with matrix inversion. This method works well as long as the scattering function is smooth and as long as one-dimensional scattering functions are being considered. However, the error introduced by the finite basis

is difficult to quantify, and no one has yet shown a tractable extension of these methods to three-dimensional scattering functions.

For these reasons, we investigated Monte Carlo methods to solve the adding equations. Monte Carlo has the advantage of being able to directly handle arbitrarily complex scattering functions and high-dimensional scattering functions. The solutions computed have no error due to approximation in a basis. There are two sub-problems to solve when applying Monte Carlo: how many terms to evaluate of the infinite sum of products of scattering functions, and how to estimate individual terms of the sum.

We solve the first problem and compute an unbiased estimate of the infinite sum by probabilistically terminating the series using Russian roulette; after computing estimates of the first few terms, we terminate with some probability after each successive term. When we continue on, subsequent terms until we do terminate are multiplied by a correction factor so that the final result is unbiased. This method was introduced to graphics by Arvo and Kirk in the framework of terminating recursive sampling for ray tracing [AK90].

The result of this termination is a finite number of terms of Equation 3.1, each of which needs to be estimated and the results summed. For each such term, we have the multiple integral represented by the composition of a set of scattering functions to estimate:

$$\int_{\mathcal{R}^+_{\mathcal{M}^2}} \cdots \int_{\mathcal{R}^+_{\mathcal{M}^2}} S_n(r' \to r_1) \cdots S_1(r_{n-1} \to r) \frac{\mathrm{d}r_{n-1}}{\mu^2_{r_{n-1}}} \cdots \frac{\mathrm{d}r_1}{\mu^2_{r_1}}$$

To compute a Monte Carlo estimate of such a term, a set of connecting rays r_i that connect the incident and outgoing rays need to be sampled (see Figure 4.7).

One approach is to start with r and sample from a distribution similar to S_1 with importance sampling to compute r_{n-1}. Given r_{n-1}, S_2 can be sampled in a similar manner, and so forth Alternatively, we can start with r' and work backwards, first sampling S_n to find r_1, etc. More generally, we can consider this to be a sampling problem analogous to sampling the path integral formulation of the equation of transfer [Vea97]; rather than always

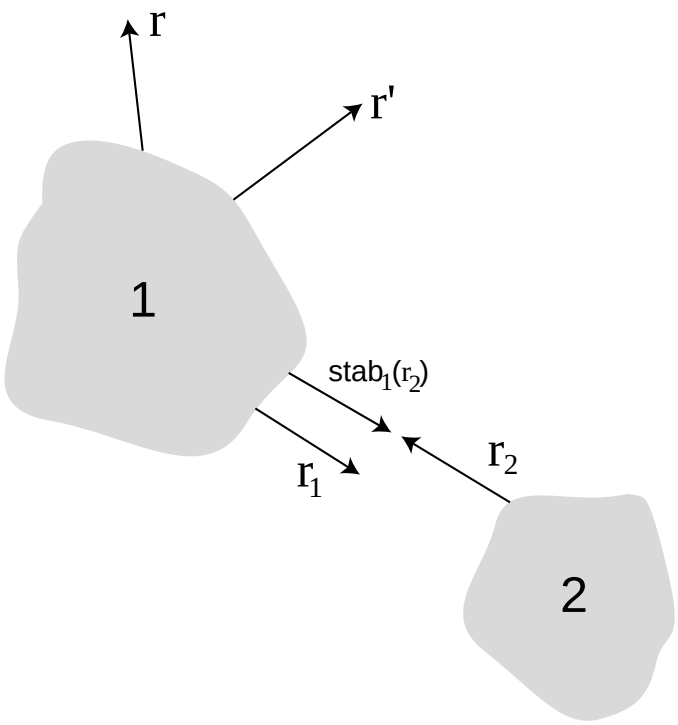


Figure 4.7: In computing the value of the aggregate scattering function $\mathbf{S}_{12}$ for the pair of rays r and r', we may need to estimate the term $\mathbf{S}_1\mathbf{S}_2\mathbf{S}_1$. In this case, we need to sample one or more pairs of connecting rays r_1 and r_2 to be able to compute an estimate of the form $\mathrm{S}_1(r \to r_1)\mathrm{S}_2(\mathrm{stab}_2(r_1) \to r_2)\mathrm{S}_1(\mathrm{stab}_1(r_2) \to r')$.

sampling rays front-to-back, back-to-front, or in some combination thereof, more flexible sampling strategies can be applied. In any case, the end result is a Monte Carlo estimate of the form

$$\frac{1}{N}\sum_{i}^{N} \frac{S_n(r' \to r_1) \cdots S_1(r_{n-1} \to r)}{p(r_1) \cdots p(r_{n-1})} \frac{1}{\mu_{r_{n-1}}^2 \cdots \mu_{r_1}^2},$$

where $p(r_i)$ gives the probability density for sampling the ray r_i. In general, this may be a conditional density based on an already-sampled ray, $p(r_i|r_{i-1})$, etc.

In sampling the intermediate rays, the operative goal is to find chains of rays such that the scattering functions S_i make a large contribution to the final result; this is analogous to the problem faced in sampling a set of rays from a light source to a sensor in classic light

transport problems.

4.4.1 Implementation

We have implemented routines that solve the adding equations in the one-dimensional setting; as such, we only need to sample connecting directions, rather than rays as in the three-dimensional case. We have implemented these routines in a modular fashion: they are given abstract data types describing the top and bottom layers as well as the incoming and outgoing directions.

The layer objects provide a small number of operations to the Monte Carlo integration routines that compute the estimates of the integrals (see Figure 4.8). In principle, the only necessary operation is evaluation of the two reflection and two transmission functions: given an incident and an outgoing angle, the fraction of incident illumination that is scattered is returned. However, there are three additional issues that must be considered.

1. In the course of choosing the intermediate connecting directions, we'd like to use importance sampling to choose these directions (see Section 3.3.3 for a discussion of general issues in importance sampling scattering functions). As such, each layer also has a routine that chooses an outgoing direction given an incident direction with importance sampling and returns the probability density of choosing that direction. For layers where distributions for importance sampling are not easily computed, a default implementation uniformly samples the hemisphere of possible directions.

2. Multiple importance sampling: once we have used a particular sampling strategy for choosing the intermediate directions (e.g. front-to-back), we can compute an estimate with lower variance with multiple importance sampling [Vea96, Vea97]. In effect, we need to compute the probability density that the other sampling strategies would have sampled the path that we did. Weights computed based on these densities are then used to scale the contribution of the sample in order to reduce variance from

`evaluate(mode, ω, ω) → value`	Evaluates R^+, R^-, T^+, or T^-, based on `mode`, for the pair of directions given.
`sample(mode, ω) → (ω′, pdf)`	Sample a direction ω' from some sampling distribution for importance sampling the scattering function. Returns the direction and the pdf for sampling the direction.
`pdf(mode, ω) → (density)`	Returns the probability density for sampling the given direction ω
`tau() → (thickness)`	Returns the optical thickness of the layer
`isDelta() → (true, false)`	Reports if a delta distribution is present

Figure 4.8: The interface provided to the adding equation estimation routines by various implementations of layers types.

importance sampling methods that pick a direction with low probability that ends up making a large contribution to the final result.

3. Delta distributions: the presence of delta distributions in the reflection and transmission functions can happen in two circumstances–either due to a surface that reflects specularly, such as a mirror, glass, or water, or due to direct transmission of attenuated light through a volumetric layer. When a delta distribution is present, the evaluation routine alone is not sufficient to describe the scattering behavior, because there is zero probability of an external sampling routine choosing a direction such that the delta distribution is non-zero. Therefore, we also include a boolean function which tells if its scattering functions are delta distribution; this is used in the estimation routine to ensure that we use the importance sampling function for this layer.

Delta distributions may be present in this series due to direct transmission through a layer of finite thickness (Equation 3.8) as well due to layers that specularly reflect or refract light (e.g. a mirror reflector at the bottom, or a Fresnel layer at the top). These are tricky because the delta functions cannot be evaluated, but only sampled—the evaluation routines always return zero. However, when such a layer samples a new direction given an incident or outgoing direction, it can pick the appropriate scattered direction. For example, when

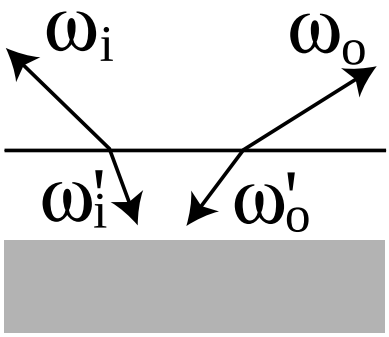


Figure 4.9: When a layer that scatters with a delta function (e.g. a Fresnel layer) is composed with another layer, the angles ω_i' and ω_o' cannot be sampled by the lower layer. Rather, the top layer must determine ω_i' directly based on ω_i (and similarly for ω_o'), since the lower layer has zero probability of randomly sampling a direction where the top layer's transmission function is non-zero.

computing the term $T^+R^+T^-$ when the top layer is a Fresnel reflector, we compute both the incident and outgoing directions to the bottom layer by sampling T^+ and T^- given the outgoing and incident directions at the top, respectively; the reflection function R^+ has no choice in sampling its incident and outgoing directions, as it would never be able to randomly find an outgoing direction that is transmitted into the final outgoing direction (see Figure 4.9).

4.4.2 Layer types

This set of operations makes it possible to implement a variety of representations for layers and easily add them together. We have implemented a number of such representations:

- A general scattering layer that takes a procedural description of the phase function and scattering and attenuation coefficients. It then uses Monte Carlo sampling to compute estimates of the reflection and transmission equations (3.7 and 3.8) for given pairs of angles.

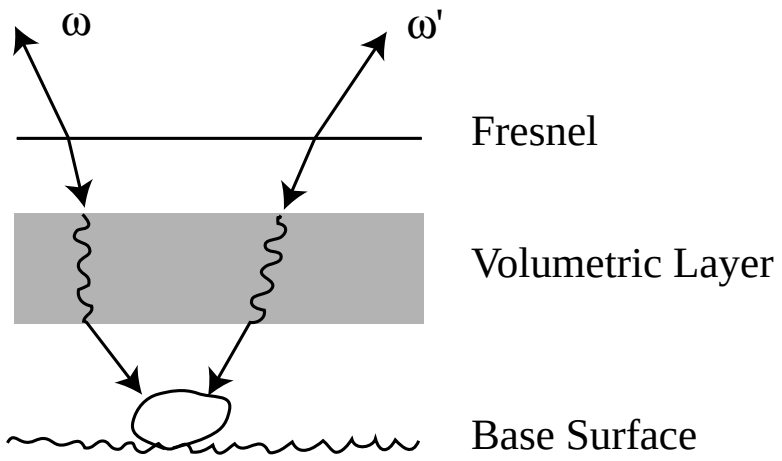


Figure 4.10: The general setting for applying the adding equation in the one-dimensional case. We have three layers to compute the aggregate scattering function of: a Fresnel reflector on top, then a volumetric scattering layer, and then a simple BRDF describing reflection at the base surface. Directions must be sampled to evaluate reflection and transmission functions between the layers.

- An opaque layer, suitable for placing at the bottom of other layers, that encapsulates a standard parameterized diffuse and glossy specular BRDF model for reflection [Hal89].

- An infinitesimally thin layer that represents scattering due to refraction at a boundary between layers with different indices of refraction. This layer accounts both for the refractive focusing and de-focusing of light as well as Fresnel effects at the boundary [HK93].

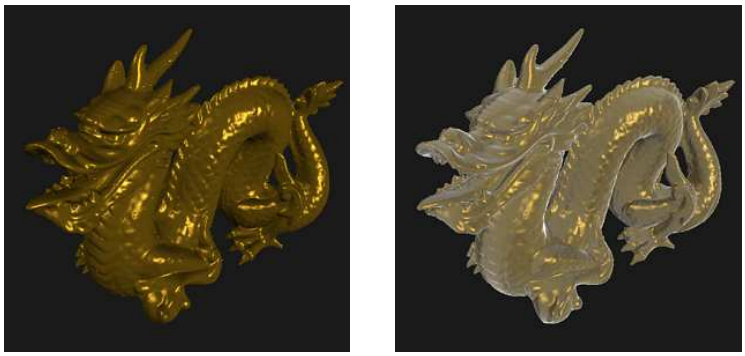
- A *thin layer approximation* that just accounts for single scattering from the phase function. As the thickness of a layer goes to zero, this approximation becomes more accurate. This representation was primarily useful for testing our implementation of the adding equations by repeatedly applying the doubling method to compute scattering from progressively thicker layers, after starting with a very thin layer.

4.5 Results

We first verified our implementation of the adding equations by comparing the scattering function values computed by applying the adding equations to two halves of an object to those computed directly for the aggregate.

To illustrate the use of the adding equations for rendering, we took a dragon model with a standard specular and diffuse surface shading model and added scattering layers, using the adding equations to compute the new scattering function that describes the composition of the base surface layer with the new scattering layer. When the routines that compute the adding equations needed to evaluate the reflection or transmission functions of the added layer, a new Monte Carlo estimate for that pair of angles was computed. With a not-very-optimized implementation, the images each took a few minutes to render on a modern PC.

The series of images in Figure 4.11 shows the results. The first image shows the object shaded with the standard shading model. As the thickness of the new layer increases going from left to right, the shiny copper base surface is gradually overwhelmed by the grey and more diffuse added layer. Eventually just a shadow of the specular highlights is left and finally no trace of the base surface once the new layer is sufficiently thick. Notice that the silhouette edges are affected more strongly by the added layer; this is because the rays traveling at oblique angles go a longer distance through the new layer. Figure 4.12 shows the result of procedurally varying the thickness of the added layer based on the local surface normal in an effort to simulate scattering effects of dust (modeled in a manner similar to Hsu and Wong [HW95]).



depth $= 0.$ depth $= 0.2$

depth $= 0.8$ depth $= 2.0$

Figure 4.11: Adding layers to a model; thicknesses are increasing from left to right. For all images, $\sigma_s = 0.5$, $\sigma_a = 0.5$, $g = -0.15$. Because the adding equations and scattering equations are used to compute the aggregate scattering function, the results accurately account for all inter-reflection inside the added layer as well as between the layer and the base surface.

Figure 4.12: Dragon model with procedurally-varying thickness of the added layer, based on local surface orientation. The result is an ad-hoc simulation of dust rendered accurately with the adding equations.

Chapter 5

Conclusion

This thesis has introduced a new theoretical framework for light scattering to computer graphics and demonstrated its application to a number of problems in rendering subsurface scattering. This theory has scattering as its basic foundation, rather than light transport, which has been the basis of almost all previous approaches to rendering. We have applied Monte Carlo integration to compute solutions to the underlying equations, extending their applicability to a wider and more general set of problems than they have previously been applied to in other fields.

For some rendering problems, results based on the scattering equation can be solved more efficiently than when the equation of transfer is used. For others, particularly those with highly scattering media, the scattering equation approach is less efficient with our current implementation. However, we believe that just as more sophisticated Monte Carlo algorithms have led to much more efficient approaches for solving the equation of transfer than the first techniques applied to it, there is similar potential to improve efficiency of solutions of the scattering equations and better take advantage of the divide-and-conquer approaches that they make possible.

The adding equations provide the foundation for correctly computing the overall scattering from collections of objects or scattering regions of space described by their individual scattering functions. They exhibit efficiencies by providing a way to break rendering problems into smaller parts and then reassemble the partial solutions; this gives a theoretical basis to clustering algorithms and a new way to apply clustering to Monte Carlo rendering algorithms. In deriving the adding equations in the ray space setting, this thesis has also introduced a unifying theory for transport problems to graphics: Preisendorfer's interaction principle. We have made connections between the interaction principle and previous approaches to rendering.

Our example of subsurface scattering as a demonstration of the three-dimensional scattering equation reflects a choice in scale rather than limitation of theory. The scattering and adding equations have applications to computing scattering from complex volumetric objects at larger scales, such as clouds, smoke, sunbeams, etc. As such, this approach has applications to the level-of-detail problem. The 3D scattering and adding equations provide the correct mathematical setting for two of the outstanding problems in level-of-detail identified by Kajiya and Kay [KK89]: automatic computation of texels from complex geometry, and computation of aggregate texels that represent two nearby texels. Furthermore, scattering functions are the correct abstraction to use to replace geometry; techniques based on BRDFs (e.g. [Kaj85, Ney98]) are inaccurate in that they do not correctly incorporate the effect of light that enters an object at a different place than it exits.

Finally, this new approach has implications for classic approaches to rendering based on light transport. By investigating the connections between this new approach and previous approaches based on light transport, it is possible to make fundamental connections between the two approaches, leading to improved algorithms in both areas.

5.1 Future Work

This theory has applications to many classic problems in rendering, including replacing geometry with scattering functions and efficiently re-rendering scenes with changes in illumination or as objects are added to or removed from them. Equally important, it has promise as a way to suggest new sampling strategies for solving the rendering equation more effectively. Understanding the connections between solution techniques that have previously been used for each of these approaches gives many directions for future work.

The recursive expansion in Monte Carlo solutions of the scattering equation solution ends up having a bidirectional sampling effect—paths are constructed in both directions and meet in the middle. Our sampling of the **SkS** term reflects a *non-local* sampling strategy [Vea97], where a scattering event at **k** is chosen before either of its adjacent scattering events have been sampled. This is in contrast to previous bidirectional sampling strategies for light transport that incrementally build paths by finding new vertices directly from a previous vertex. As such, understanding the connections between the path sampling strategies that we have used and previous bidirectional path sampling strategies is important future work. Furthermore, techniques that ameliorate the exponential nature of the recursive sampling and more effectively re-use sub-paths should improve performance in cases where the albedo is high. Another area for further investigation is better importance sampling techniques for the 3D case and the application of multiple importance sampling to reduce variance.

5.1.1 Geometric Scenes

Extending this approach to rendering geometric scenes should be straightforward; the basic building blocks of the algorithm (computing the zero bounce, one bounce, and multiple bounce scattering) of regions of space can be easily implemented for scenes comprised of geometric objects. The only immediate complication is that all scattering must happen at

surfaces, rather than at more arbitrarily-chosen points in space. However, we do not believe that this is a substantial difficulty, as it is effectively the introduction of delta distributions to the definitions of the scattering functions involved, thus reducing the number of dimensions to be considered in the integrals.

5.1.2 Approximation and Biased Approaches

Considering different approaches for computing or approximating the scattering functions of regions of space should be fruitful. In particular, biased Monte Carlo techniques or techniques that approximate the eight-dimensional scattering function with lower-dimensional approximations may be quite efficient. Because interactions between regions of space would still be considered accurately, the results may be sufficiently accurate for some applications. In particular, given an approximation to the scattering function of a region $\tilde{S}$, given a bound of the error with the actual scattering function $|\tilde{S} - S|$, it may be possible to bound the error in the overall result after computing the interactions.

Another interesting potential application is to the level of detail problem. If approximate scattering functions of unimportant regions of space can be quickly computed (e.g. using a lower-resolution geometric model), but only when these regions have a low contribution to the overall scattering function, accurate approximations should be computed. We also note that Neyret's ellipsoidal normal distribution functions [Ney95, Ney98] would be a useful way to represent scattering functions of collections of geometry.

5.1.3 Hierarchical Scattering

Hierarchical algorithms have previously been very effective for radiosity methods in computer graphics [HSA91]; by taking advantage of the form of interactions in the radiosity matrix, they are able to compute solutions efficiently by carefully choosing the level of refinement at which each pair of objects interacts. The adding equations give the appropriate

theoretical basis for developing hierarchical algorithms for rendering arbitrary scenes, not just those under the radiosity assumptions.

Specifically, a 3D scene can be hierarchically decomposed into smaller volumes of space, for example using an octree or a kd-tree. The overall scattering function of the scene can then be written directly in terms of integrals of the scattering functions of the smaller regions, accounting for all possible scattering interactions between these regions. By using a variety of methods to estimate the scattering functions of individual regions and assembling these sub-solutions appropriately, there may be opportunities to develop new algorithms for efficiently rendering complex geometric environments. In particular, if the regions at the leaves of the tree were small enough to be effectively homogeneous, the dipole approach introduced to graphics by Jensen et al [JMLH01, JB02] could be an effective way to efficiently compute the scattering functions at the leaves.

One advantage of such an approach is that computation is naturally localized in the regions of the scene where most of the scattering is occurring; little effort is spent on areas that make a small contribution to the final result. Further, there are opportunities to re-use computation across multiple computations of scattering function values, further improving efficiency [Esp79].

Finally, with this approach there is an opportunity to precompute scattering functions for some regions of space, and using those to approximate the scattering function $\mathbf{S}_x$, rather than evaluating it directly with Monte Carlo integration. If accurate approximations can be developed and the approximations are used only when they make a small contribution to the overall answer, there is further potential for efficiency improvements.

Bibliography

[AK90] James Arvo and David Kirk. Particle transport and image synthesis. *Computer Graphics*, 24(4):63–66, August 1990.

[Amb42] V. A. Ambarzumian. A new method for computing light scattering in turbid media. *Izv. Akad. Nauk SSSR*, 3, 1942.

[Amb43] V. A. Ambarzumian. Diffuse reflection of light by a foggy medium. *Comtes Rendus (Doklady) de L'Académie des Sciences de l'URSS*, 38(8):229–232, 1943.

[Amb58] V. A. Ambarzumian, editor. *Theoretical Astrophysics*. Pergamon Press, New York, New York, 1958.

[Arv93a] James Arvo. Linear operators and integral equations in global illumination. In *Global Illumination, SIGGRAPH '93 Course Notes*, volume 42, August 1993.

[Arv93b] James Arvo. Transfer equations in global illumination. In *Global Illumination, SIGGRAPH '93 Course Notes*, volume 42, August 1993.

[Arv95] James Arvo. *Analytic Methods for Simulated Light Transport*. PhD thesis, Yale University, December 1995.

[Atk93] Kendall Atkinson. *Elementary Numerical Analysis*. John Wiley & Sons, New York, 1993.

[BK56] Richard Bellman and Robert Kalaba. On the principle of invairant imbedding and propagation through inhomoheneous media. *Proceedings of the National Academy of Sciences*, 42:629–632, 1956.

[BKP63] Richard E. Bellman, Robert E. Kalaba, and Marcia C. Prestrud. *Invariant Imbedding and Radiative Transfer in Slabs of Finite Thickness*. American Elsevier Publishing Company, New York, 1963.

[BKU62] Richard Bellman, Robert Kalaba, and Sueo Ueno. Invariant imbedding and diffuse reflection from a two-dimensional flat layer. *Icarus*, 1:297–303, 1962.

[BKW60] Richard Bellman, Robert Kalaba, and G. Milton Wing. Invariant imbedding and the reduction of two point boundary value problems to initial value problems. *Proceedings of the National Academy of Sciences*, 46:1646–1649, 1960.

[Bli82] James F. Blinn. Light reflection functions for simulation of clouds and dusty surfaces. *Computer Graphics*, 16(3):21–29, July 1982.

[Bus61] Ida Busbridge. On inhomogeneous stellar atmospheres. *Astrophysical Journal*, 133:198–209, 1961.

[BW75] Richard E. Bellman and G. M. Wing. *An Introduction to Invariant Imbedding*. John Wiley & Sons, New York, 1975.

[Cha60] S. Chandrasekar. *Radiative Transfer*. Dover Publications, New York, 1960. Originally published by Oxford University Press, 1950.

[Cor89] James Cornones. A discrete model of transport and invariant imbedding in two or more dimensions. In Paul Nelson, V. Farber, Thomas A. Manteuffel, Daniel L. Seth, and Jr. Andrew B. White, editors, *Transport Theory, Invariant Imbedding, and Integral Equations*, pages 271–278. Marcel Dekker, Inc., New York, 1989.

[CW93] Michael F. Cohen and John R. Wallace. *Radiosity and Realistic Image Synthesis*. Academic Press Professional, San Diego, CA, 1993.

[DEL+99] Julie Dorsey, Alan Edelman, Justin Legakis, Henrik Wann Jensen, and Hans Køhling Pedersen. Modeling and rendering of weathered stone. *Proceedings of SIGGRAPH 99*, pages 225–234, August 1999.

[DHT+00] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 145–156, July 2000.

[Esp79] Larry W. Esposito. An ‘adding’ algorithm for the markov chain formalism for radiation transfer. *Astrophysical Journal*, 233:661–663, 1979.

[Fan81] Ronald L. Fante. Relationship between radiative-transport theory and Maxwell’s equations in dielectric media. *Journal of the Optical Society of America*, 71(4):460–468, April 1981.

[Fis96] George S. Fishman. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer-Verlag, New York, 1996.

[FSW89] V. Faber, Deniel L. Seth, and G. Milton Wing. Invariant imbedding in two dimensions. In Paul Nelson, V. Farber, Thomas A. Manteuffel, Daniel L. Seth, and Jr. Andrew B. White, editors, *Transport Theory, Invariant Imbedding,*

and Integral Equations, pages 279–300. Marcel Dekker, Inc., New York, 1989.

[Ger39] A. Gershun. The light field. *Journal of Mathematics and Physics*, 18:51–151, 1939.

[Gla95] Andrew Glassner. *Principles of Digital Image Synthesis*. Morgan Kaufmann, New York, 1995.

[GY89] R. M. Goody and Y. L. Yung. *Atmospheric Radiation*. Oxford University Press, 1989.

[Hal89] Roy Hall. *Illumination and Color in Computer Generated Imagery*. Springer-Verlag, New York, 1989. includes C code for radiosity algorithms.

[Han71] James E. Hansen. Multiple scattering of polarized light in planetary atmospheres. Part I. The doubling method. *Journal of the Atmospheric Sciences*, 28:120–125, January 1971.

[HG41] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *Astrophysical Journal*, 93:70–83, 1941.

[HK93] Pat Hanrahan and Wolfgang Krueger. Reflection from layered surfaces due to subsurface scattering. In *Computer Graphics* Proceedings, pages 165–174, August 1993.

[HM92] Chet S. Haase and Gary W. Meyer. Modeling pigmented materials for realistic image synthesis. *ACM Transactions on Graphics*, 11(4):305–335, October 1992.

[Hov69] J. W. Hovenier. Symmetry relationships for scattering of polarized light in a slab of randomly oriented particles. *Journal of the Atmospheric Sciences*, 26:488–499, 1969.

[Hov77] J. W. Hovenier. A unified treatment of the reflected and transmitted intensities of a homogeneous plane-parallel atmosphere. *Astronomy and Astrophysics*, 68:239–250, 1977.

[HSA91] Pat Hanrahan, David Salzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. In Thomas W. Sederberg, editor, *Computer Graphics (SIGGRAPH '91 Proceedings)*, volume 25, pages 197–206, July 1991.

[HT74] James E. Hansen and Larry D. Travis. Light scattering in planetary atmospheres. *Space Science Reviews*, 16:527–610, 1974.

[HTSG91] Xiao D. He, Kenneth E. Torrance, Francois X. Sillion, and Donald P. Greenberg. A comprehensive physical model for light reflection. In Thomas W. Sederberg, editor, *Computer Graphics (SIGGRAPH '91 Proceedings)*, volume 25, pages 175–186, July 1991.

[HW95] Siu-Chi Hsu and Tien-Tsin Wong. Simulating dust accumulation. *IEEE Computer Graphics and Applications*, 15(1):18–25, January 1995.

[Irv75] William M. Irvine. Multiple scattering in planetary atmospheres. *Icarus*, 25:175–204, 1975.

[JB02] Henrik Wann Jensen and Juan Buhler. A rapid hierarchical rendering technique for translucent materials. *ACM Transactions on Graphics*, 21(3):576–581, July 2002.

[JMLH01] Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. In *Proceedings of ACM SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, pages 511–518, August 2001.

[Kaj85] James T. Kajiya. Anisotropic reflection models. In *Computer Graphics (SIGGRAPH '85 Proceedings)*, volume 19, pages 15–21, July 1985.

[Kaj86] James T. Kajiya. The rendering equation. *Computer Graphics*, 20(4):143–150, August 1986.

[KK89] James T. Kajiya and Timothy L. Kay. Rendering fur with three dimensional textures. *Computer Graphics*, 23(3):271–280, July 1989.

[KM31] P. Kubelka and F. Munk. Ein Beitrag zur Optik der Farbanstriche. *Z. Tech. Physik.*, 12:593, 1931.

[Kou69] V. Kourganoff. *Introduction to the General Theory of Particle Transfer*. Gordon and Breach, New York, 1969.

[KP68] George W. Kattawar and Gilbert N. Plass. Radiance and polarization of multiple scattered light from haze and clouds. *Applied Optics*, 7(8):1519–1527, 1968.

[Kru90] Wolfgang Krueger. Volume rendering and data feature enhancement. 24(5):21–26, November 1990.

[KV84] James T. Kajiya and Brian P. Von Herzen. Ray tracing volume densities. *Computer Graphics*, 18(3):165–173, July 1984.

[KW86] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods: Volume I: Basics*. John Wiley & Sons, New York, 1986.

[Laf96] Eric Lafortune. *Mathematical Models and Monte Carlo Algorithms for Physically Based Rendering*. PhD thesis, Katholieke Universiteit Leuven, February 1996.

[LM80] Jong-Sen Lee and R. R. Meier. Angle-dependent frequency redistribution in a plane-parallel medium: External source case. *Astrophysical Journal*, 240:185–195, 1980.

[LW94] Eric Lafortune and Yves Willems. A theoretical framework for physically based rendering. *Computer Graphics Forum*, 13(2):97–107, 1994.

[McC05] J. A. McClelland. On secondary radiation (part II) and atomic structure. *Scientific Transactions of the Royal Dublin Society*, 9(1):1–8, 1905.

[McC06] J. A. McClelland. The energy of secondary radiation. *Scientific Transactions of the Royal Dublin Society*, 9(1):9–26, 1906.

[McC94] William Ross McCluney. *Introduction to radiometry and photometry*. Artech House, 1994.

[MH06a] J. A. McClelland and F. E. Hackett. The absorption of β radium rays by matter. *Scientific Transactions of the Royal Dublin Society*, 9(1):37–50, 1906.

[MH06b] J. A. McClelland and F. E. Hackett. Secondary radiation from compounds. *Scientific Transactions of the Royal Dublin Society*, 9(1):27–36, 1906.

[ML78] R. R. Meier and Jong-Sen Lee. A Monte Carlo study of frequency redistribution in an externally excited medium. *Astrophysical Journal*, 219:262–273, January 1978.

[ML81] R. R. Meier and Jong-Sen Lee. Angle-dependent frequency redistribution: Internal source case. *Astrophysical Journal*, 250:376–383, 1981.

[MM98] Gavin Miller and Marc Mondesir. Rendering hyper-sprites in real time. *Eurographics Rendering Workshop 1998*, pages 193–198, June 1998.

[MMKW97] Nelson Max, Curtis Mobley, Brett Keating, and En-Hua Wu. Plane-parallel radiance transport for global illumination in vegetation. In *Eurographics Rendering Workshop 1997*, pages 239–250. Eurographics, Springer Wien, June 1997.

[Ney95] Fabrice Neyret. A general and multiscale model for volumetric textures. In Wayne A. Davis and Przemyslaw Prusinkiewicz, editors, *Graphics Interface '95*, pages 83–91. Canadian Human-Computer Communications Society, May 1995.

[Ney98] Fabrice Neyret. Modeling, animating, and rendering complex scenes using volumetric textures. *IEEE Transactions on Visualization and Computer Graphics*, 4(1), January – March 1998.

[NRH+77] Fred E. Nicodemus, J. C. Richmond, J. J. Hisa, I. W. Ginsberg, and T. Limperis. *Geometrical Considerations and Nomenclature for Reflectance*. Monograph number 160. National Bureau of Standards, Washington DC, 1977.

[NS97] Paul Nelson and Daniel L. Seth. Integrodifferential equations for the two-dimensional transition kernels of invariant imbedding. *Applied Mathematics and Computation*, 82:67–83, 1997.

[NSV92] Paul Nelson, D. L. Seth, and R. Vasudevan. An integrodifferential equation for the two-dimensional reflection kernel. *Applied Mathematics and Computation*, 49:1–18, 1992.

[NUW98] H. H. Natsuyama, S. Ueno, and A. P. Wang. *Terrestial Radiative Transfer*. Springer-Verlag, Hong Kong, 1998.

[Per85] Ken Perlin. An image synthesizer. In *Computer Graphics (SIGGRAPH '85 Proceedings)*, volume 19, pages 287–296, July 1985.

[Per99] A. Peraiah. Principles of invariance in radiative transfer. *Space Science Reviews*, 87:465–538, 1999.

[PH00] Matt Pharr and Pat Hanrahan. Monte Carlo evaluation of non-linear scattering equations for subsurface reflection. Computer Graphics Proceedings, Annual Conference Series, pages 75–84. ACM Press / ACM SIGGRAPH / Addison Wesley Longman, July 2000.

[PK68] Gilbert N. Plass and George N. Kattawar. Monte Carlo calculations of light scattering from clouds. *Applied Optics*, 7(3):415–419, 1968.

[PP51] Glenn H. Peebles and Milton S. Plesset. Transmission of gamma-rays through large thicknesses of heavy materials. *Physical Review*, 81(3):430–439, 1951.

[Pre58a] Rudolph W. Preisendorfer. Function relations for the R and T operators on plane-parallel media. *Proceedings of the National Academy of Sciences*, 44:323–327, 1958.

[Pre58b] Rudolph W. Preisendorfer. Invariant imbedding relation for the principles of invariance. *Proceedings of the National Academy of Sciences*, 44:320–323, 1958.

[Pre65] Rudolph W. Preisendorfer. *Radiative Transfer on Discrete Spaces*. Pergamon Press, Oxford, 1965.

[Pre76] R. W. Preisendorfer. *Hydrologic Optics*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, Honolulu, Hawaii, 1976. Six volumes.

[Red62] Raymond Redheffer. On the relation of transmission-line theory to scattering and transfer. *Journal of Mathematics and Physics*, 41:1–41, 1962.

[RPV93] Holly Rushmeier, Charles Patterson, and Aravindan Veerasamy. Geometric simplification for indirect illumination calculations. In *Proceedings of Graphics Interface '93*, pages 227–236, May 1993.

[Rus88] Holly E. Rushmeier. *Realistic Image Synthesis for Scenes with Radiatively Participating Media*. Ph.d. thesis, Cornell University, 1988.

[Sch05] Arthur Schuster. Radiation through a foggy atmosphere. *Astrophysical Journal*, 21(1):1–22, January 1905.

[SD95] François Sillion and George Drettakis. Feature-based control of visibility error: A multi-resolution clustering algorithm for global illumination. In *SIGGRAPH 95 Conference Proceedings*, pages 145–152. Addison Wesley, August 1995.

[SDS95] François Sillion, G. Drettakis, and Cyril Soler. A clustering algorithm for radiance calculation in general environments. In *Eurographics Rendering Workshop 1995*. Eurographics, June 1995.

[SG69] Jerome Spanier and Ely M. Gelbard. *Monte Carlo principles and neutron transport problems*. Addison-Wesley, Reading, Massachusetts, 1969.

[Sta99] Jos Stam. Diffraction shaders. *Proceedings of SIGGRAPH 99*, pages 101–110, August 1999.

[Sto62] George Stokes. On the intensity of the light reflected from or transmitted through a pile of plates. *Proceedings of the Royal Society*, January 1862. Reprinted in Mathematical and Physical Papers of Sir George Stokes, Volume IV, Cambridge, 1904.

[Tes87] J. Tessendorf. Radiative transfer as a sum over paths. *Physical Review A: General Physics*, 35(2):872–878, January 1987.

[Tes90] J. Tessendorf. Radiative transfer on curved surfaces. *Journal of Mathematical Physics*, 31(4):1010–1019, 1990.

[TJH66] S. Twomey, H. Jacobowitz, and H. B. Howell. Matrix methods for multiple-scattering problems. *Journal of the Atmospheric Sciences*, 23:289–296, May 1966.

[TL80] A. E. Taylor and D. C. Lay. *Introduction to functional analysis*. John Wiley & Sons, New York, 1980.

[TS67] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9), 1967.

[Uen60] Sueo Ueno. The probabilistic method for problems of radiative transfer. X. Diffuse reflection and transmission in a finite inhomogeneous atmosphere. *Astrophysical Journal*, 132:729–745, 1960.

[van71] H. C. van de Hulst. Multiple scattering in planetary atmospheres. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 11:785–795, 1971.

[van80] Hendrik Christoffel van de Hulst. *Multiple Light Scattering*. Academic Press, New York, 1980. Two volumes.

[van81] Hendrik Christoffel van de Hulst. *Light Scattering by Small Particles*. Dover Publications, New York, 1981. Originally published by John Wiley and Sons, 1957.

[Vea96] Eric Veach. Non-symmetric scattering in light transport algorithms. In Xavier Pueyo and Peter Schröder, editors, *Eurographics Rendering Workshop 1996*. Springer Wien, June 1996.

[Vea97] Eric Veach. *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, December 1997.

[VG94] Eric Veach and Leonidas Guibas. Bidirectional estimators for light transport. In *Fifth Eurographics Workshop on Rendering*, pages 147–162, Darmstadt, Germany, June 1994.

[VG95] Eric Veach and Leonidas J. Guibas. Optimally combining sampling techniques for Monte Carlo rendering. In *Computer Graphics* Proceedings, pages 419–428, August 1995.

[Wan90] Alan P. Wang. Basic equations of three-dimensional radiative transfer. *Journal of Mathematical Physics*, 31(1):175–181, January 1990.

[Wan93] Alan P. Wang. Searchlight on a target with diffuse background. I. *Journal of Mathematical Physics*, 34(2):878–884, February 1993.

[WAT92] Stephen Westin, James Arvo, and Kenneth Torrance. Predicting reflectance functions from complex surfaces. *Computer Graphics*, 26(2):255–264, July 1992.

[WU94] A. P. Wang and S. Ueno. Searchlight on a target with diffuse background. II. *Computers and Mathematics, with Applications*, 27:169–174, 1994.

[Yan97] Edgard G. Yanovitskij. *Light Scattering In Inhomogeneous Atmospheres*. Springer-Verlag, 1997.

[ZWCS99] Douglas E. Zongker, Dawn M. Werner, Brian Curless, and David H. Salesin. Environment matting and compositing. In *SIGGRAPH 99 Conference Proceedings*, pages 205–214. Addison Wesley, August 1999.

Index