

Transfer Equations in Global Illumination*

James Arvo
Program of Computer Graphics
Cornell University

Abstract

The purpose of these notes is to describe some of the physical and mathematical properties of the equations occurring in global illumination. We first examine the physical assumptions that make the particle model of light an appropriate paradigm for computer graphics and then derive a balance equation for photons. In doing this we establish connections with the field of *radiative transfer* and its more abstract counterpart, *transport theory*. The resulting balance equation, known as the *equation of transfer*, accounts for large-scale interaction of light with participating media as well as complex reflecting surfaces. Under various simplifying assumptions the equation of transfer reduces to more conventional equations encountered in global illumination.

1 Introduction

Global illumination connotes a physically-based simulation of light appropriate for synthetic image generation. The task of such a simulation is to model the interplay of light among large-scale objects of an environment in order to approximate the quantity and quality of light reaching the eye of an observer. The problem is *global* in that all objects can interact with one another by means of scattered light. To properly simulate the lighting of a room, for instance, the entire surroundings must be taken into account; architectural features and even the interior decoration can affect the overall illumination [32].

But why the emphasis on physically-based simulation? The answer is that we wish to accurately predict the appearance of a hypothetical scene, and by imitating the physics we might thereby approximate the visual stimulus of viewing the scene. Of course, there is more to it than physics. In addition to the behavior of light, there are many important considerations in producing a convincing image; for instance, the characteristics of the display device, the physiology of the eye, and even the psychology of perception.

Although global illumination does not reduce to physics and computation alone, these aspects currently dominate the field. Consequently, global illumination is predominantly the study of algorithms for simulating the interaction of visible light with large-scale geometry; that is, geometry with at least some features significantly larger than the wavelength of the light. These goals are by no means unique to global illumination.

Global illumination can be considered a subfield of *radiative transfer*, which is the study of the interaction of radiant energy with matter at the macroscopic scale [7, 43]. The origin

*Copyright ©1993 by James Arvo

of radiative transfer theory is generally attributed to the astrophysical work of Schuster [48] and Schwarzschild [49] near the turn of the century. Since then, it has found application in areas such as illumination engineering [31], radiative heat transfer [51], hydrologic optics [44], agriculture [34, 33], remote sensing [3, 12], vision [64], and computer graphics [13]. The central tenets of radiative transfer are embodied in the *equation of transfer*; an integro-differential equation describing the gross interaction of light with participating media and, given appropriate boundary conditions, its interaction with arbitrary surfaces. Although rendering methods exist for simulating participating media [21, 46], the origin and extensive history of the equation of transfer are not widely recognized within the computer graphics community. Studying the equation of transfer is an excellent way to build a bridge to the vast literature in radiative transfer and its direct descendants, and this is one purpose of these notes.

Global illumination can also be viewed as a subfield of *transport theory* [2, 17, 9], a field that encompasses all macroscopic phenomena resulting from the interaction of infinitesimal particles with a medium. The macroscopic behaviors of photons, neutrons, and gas molecules are all within its purview. The central equation of transport theory is known as the *Boltzmann equation*.

Placing global illumination in the larger context of transport theory reveals common ground with fields such as neutron transport theory and the kinetic theory of gases. There are advantages in exploring the similarities. While neutron migration may appear to have little to do with rendering a foggy atmosphere, the equations governing the two phenomena are virtually identical. Consequently, we can gain insight from the older, more developed theories that describe such phenomena. As a more practical matter, many numerical techniques have been devised in related fields, and these are only beginning to be explored in global illumination. A wide assortment of Monte Carlo techniques have been developed for neutral particle transport [27, 55] that are relevant to global illumination [1]. The concept of *importance* for view-dependent radiosity [54] is another example of a technique adapted from a related field [26].

In the remainder of this section we examine some of the the assumptions of global illumination and characterize the underlying physical model. In section 2 we derive a well-known transport equation for abstract particles. In section 3 we specialize the abstract equation to radiative transfer and deduce from it various equations occurring in computer graphics.

1.1 The Physical Basis of Global Illumination

What does it mean for global illumination to be physically-based? Certainly it implies that the simulation of light is in some sense faithful to the actual physical behavior of light. Yet no image synthesis technique in use today captures the full repertoire of light. Ray tracing [61] and radiosity [13, 38] feature nearly mutually exclusive models of light reflection: each valid in itself, but incomplete. Methods that combine ray tracing and radiosity typically neglect more complex modes of reflection [53]. The rendering equation [19] and methods that incorporate complex surface reflection [4, 52] still neglect light scattering by participating media such as smoke. Methods that model participating media have thus far limited the forms of surface reflection [46, 47, 21]. Finally, absent from *all* of the cited approaches are macroscopic effects due to interference and diffraction. Every global illumination algorithm assumes some model of light and, with the possible exception of quantum

electrodynamics [10], every model of light is itself incomplete.

Maxwell's equations provide a very detailed description of light, accounting for a wide variety of phenomena including dispersion, interference, and diffraction. However, this *physical optics* level of description is overly detailed for global illumination. It is most relevant at scales on the order of the wavelength of light, where wave optics effects predominate. Geometrical optics, on the other hand, deals with the large-scale behavior of light, ignoring the wave-related details of physical optics. However it leaves out too much. Pure geometrical optics is insufficient for global illumination because it incorporates no *radiometric* concepts for the measure of light [22].

For global illumination, we need to model visual phenomena; those that are observable by the eye or any instrument capable of discriminating light intensity and frequency. This is called a *phenomenological* level of description as it focuses on phenomena corresponding to percepts such as brightness and color. Because the principles of radiative transfer operate at a size scale large compared to the wavelength of light, and a time scale large compared to the frequency, it is essentially a phenomenological description of the behavior of light [43]. Thus, radiative transfer matches the needs of global illumination.

Radiative transfer does not provide explanations of phenomena at the level of quantum mechanics or electromagnetism, yet it must incorporate information from these more detailed levels. Theories at the microscopic scale are usually needed to predict first-order effects such as local scattering and absorption, which then enter into the simulations at the macroscopic scale. For instance, the physically-based reflection model of He [16] employs wave optics to characterize reflection from rough surfaces, and this model can be used for global illumination [52]. For many materials with complex microgeometries, there exists a natural hierarchy of scales, with geometrical optics taking over at the point when wave effects become negligible [60]. Another avenue by which physical optics effects can enter into radiative transfer is through physical measurements; incorporating bidirectional reflectance functions of actual materials is a prime example [57].

1.2 Dimensions and Units

Every physically-based simulation ultimately concerns physical *quantities*; that is, measurable attributes of phenomena or of matter [35]. Implicit in any equation expressing a relationship among physical quantities is a system of units in which the quantities were measured. Thus, each algebraic equation with a physical interpretation has a corresponding *dimensional equation* showing how the units combine [18]. Unfortunately, the units for radiometric quantities can be troublesome, owing to the overwhelming number of distinct and sometimes inconsistent definitions [30, 56, 36, 37, 20]. As we begin to consider domains outside of radiative transfer the problem grows worse; each field has its own bewildering assortment of units and terminology. This unfortunate fact is largely historical, resulting from independent development of many fundamentally related fields.

One defense against the onslaught of unfamiliar units and terminology is dimensional analysis; that is, explicit examination of the associated dimensional equations. While dimensional analysis cannot determine unknown coefficients, it can reveal necessary relationships among the quantities involved [18], and provides a first check of validity. Throughout these notes, physical units will be discussed along with each quantity. We shall use the standard notational device of brackets to the right of an equation to indicate the units possessed by each term of the expression or equation.

The present discussion is restricted to the physical domain, so we need not introduce luminous intensity or its corresponding unit, the candela. This will keep the proliferation of units to a minimum. The only units that we shall have occasion to use are the following, from the international system of units [35]:

<i>Quantity</i>	<i>Units</i>	<i>Symbol</i>
time	second	s
length	meter	m
solid angle	steradian	sr
energy	joule	J
power	watt	W

Time, length, and solid angle are base units, while energy and power are derived units. The joule has base units of $\text{m}^2 \text{kg s}^{-2}$, where kg is the unit of mass, and the watt is joules per second. For our purposes we needn't reduce energy to anything more fundamental. We can simply treat the joule as a base unit corresponding to some irreducible quantity that nature has provided.

2 A Balance Equation for Particles

The macroscopic phenomena modeled by radiative transfer are accounted for by the particle model of light. Consequently, radiative transfer is synonymous with *photon transport*, a term connoting large-scale statistical behaviors of photons. This perspective leads naturally to Monte Carlo simulations, in which large numbers of photon histories are tallied [50], but it also provides a convenient conceptual model.

In this section we approach radiative transfer by first reducing the model of light to the bare essentials, leaving behind only abstract particles. This captures the spirit of *transport theory*, a theory whose goal is to determine the distribution of abstract particles in space and time, taking into account their motion and interaction with a host medium. The resulting distribution is expressed purely in terms of the geometrical and physical properties of the medium [5, 9].

Transport theory is classical in nature, yet its focus on discrete particles differentiates it from classical theories such as electromagnetism. It requires many assumptions concerning the nature of these particles [27, 62]. The two most important assumptions are that 1) the particles are so small and numerous that their statistical distribution can be treated as a continuum, and 2) at any point in time a particle is completely characterized by its position and velocity, and internal states such as polarization, frequency, charge, or spin. These assumptions lead naturally to the notion of *phase space*.

Phase space is an abstraction for dealing with configurations of particles. Each dimension of phase space represents a degree of freedom of a particle; commonly six in all. For simplicity, we shall only consider particles that move with a constant speed v and possess no internal states. This restricted problem corresponds to *gray* or *mono-energetic* radiative transfer, or *one-speed* transport for generic particles. The one-speed assumption requires only five degrees of freedom per particle; three for position and two for direction. The corresponding 5-dimensional phase space is $\mathcal{R}^3 \times \mathcal{S}^2$ where \mathcal{R}^3 is Euclidian 3-space and \mathcal{S}^2 is the unit sphere in \mathcal{R}^3 .

To exactly characterize a configuration of particles we must specify all degrees of freedom of each particle at an instant in time; this corresponds to discrete points in phase space. Such a complete accounting for discrete particles has a natural analogue for a continuum of particles; namely, *phase space density*. Phase space density, denoted by $n(\mathbf{r}, \omega, t)$, is a real-valued function defined over phase space and time such that

$$n(\mathbf{r}, \omega, t) d\mathbf{r} d\omega \quad (1)$$

is the number of particles in a differential volume $d\mathbf{r}$ about the point $\mathbf{r} \in \mathcal{R}^3$ moving in a direction within a differential solid angle $d\omega$ about $\omega \in \mathcal{S}^2$ at time t . Thus, its units are $\text{m}^{-3}\text{sr}^{-1}$. Phase space density makes no mention of material attributes such as mass or energy; these concepts do not enter until we assign physical meaning to the particles, as we do in section 3. The abstract nature of phase space density makes it quite universal, it underlies virtually all particle transport problems from radiative transfer to gas dynamics.

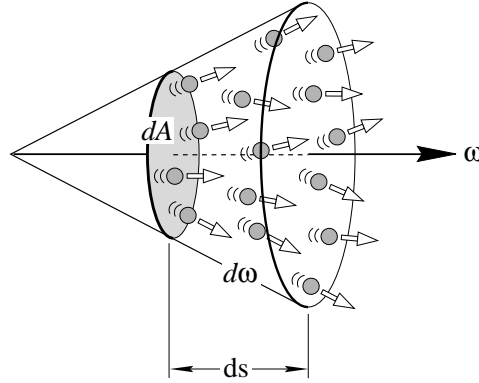


Figure 1: *Phase space flux is the number of particles crossing the surface dA perpendicular to ω per unit area, per unit solid angle, per unit time.*

It is frequently more convenient to characterize the density of particles by their rate of flow across a real or imaginary surface. Consider the particles that pass through a differential area dA with directions in differential solid angle $d\omega$ about the surface normal in time dt . All the particles are contained within the volume $dA ds$, where $ds = v dt$, as shown in Figure (1). Assuming that \mathbf{r} is a point within this volume, the number of particles is

$$n(\mathbf{r}, \omega) dA ds d\omega. \quad (2)$$

But rather than focusing on the particles within a volume can shift emphasis to the surface dA and the rate at which particles cross it by defining *phase space flux* ϕ :

$$\phi(\mathbf{r}, \omega) \equiv v n(\mathbf{r}, \omega) \quad \left[\frac{1}{\text{m}^2 \text{sr s}} \right]. \quad (3)$$

In terms of phase space flux, the number of particles in $dA ds$ is

$$\phi(\mathbf{r}, \omega) dA d\omega dt. \quad (4)$$

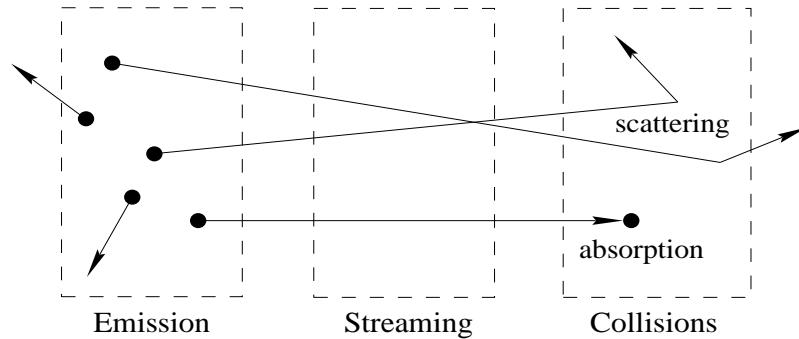


Figure 2: *Particle behaviors fall into three categories: emission, streaming, and collisions. The latter is further divided into absorption and scattering.*

Phase space flux will be the fundamental quantity of interest; it can be used to express virtually any quantity relating to a distribution of particles. Our goal is to derive an equation for the function ϕ based on particle conservation; this requires that we characterize all properties of the particles that can affect their motion through the medium. We observe that all particle behaviors fall into one of three categories: *emission*, *streaming*, and *collisions*.

Emission is the name given to any physical process that injects new particles into the system. Physically, emission may result from chemical, thermal, or even nuclear processes. Once emitted, particles travel until they suffer an instantaneous collision. For neutral particles, such as photons and neutrons, the path followed is independent of external forces, and the behavior of following such a path is called streaming. For photons this path is determined by the *eiconal equation* [28] which accounts for continuously varying index of refraction. If we neglect refraction, as we shall do here, streaming is in straight lines.

Collisions imply interactions with matter, and can be further subdivided into *absorption* and *scattering*. Absorption removes a particle from the system by capturing it and converting its energy into some other form, such as thermal motion of the medium. Scattering is a process that instantaneously changes the direction of travel of a particle. See Figure (2). If the speed of travel or energy of the particle remains unchanged by the scattering it is called *elastic*. All scattering is assumed to be elastic in one-speed transport.

A very common assumption concerning collisions is that particles do not interact with one another; or more precisely, that such interactions are negligible. In the absence of this assumption, the rate of collision is itself a function of the particle distribution; a situation leading to a non-linear balance equation. This source of non-linearity is sometimes unavoidable, as in some problems of gas dynamics [14]. To preserve linearity in photon transport it is necessary to ignore effects such as interference, which imply photon-photon interactions. In neutron transport, where inter-particle collisions are exceedingly rare, this assumption is almost universally made.

2.1 An Abstract Transport Equation

In this section we derive a broadly applicable balance equation for neutral non-interacting particles; the same equation appears in both radiative transfer and neutron transport theory. The approach we shall take is similar to that of Duderstadt and Martin [9], and is sometimes

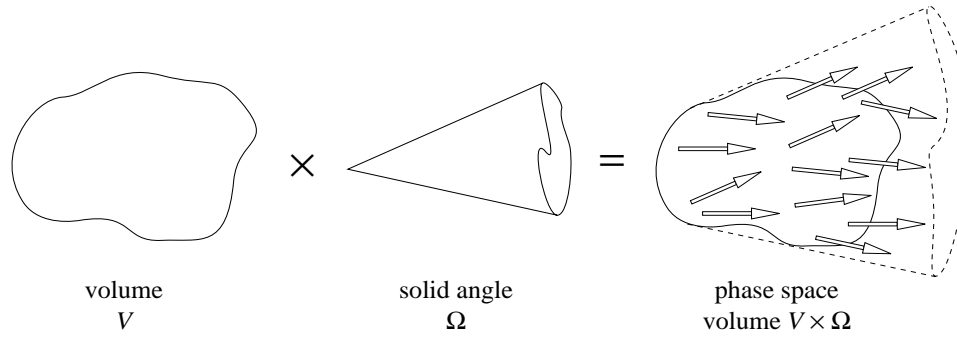


Figure 3: *The Cartesian product of a volume and a solid angle results in a five-dimensional volume in phase space.*

called the *Eulerian* derivation [40] since the analysis is carried out with respect to a fixed reference frame.

To derive a balance equation for the particles, we begin by examining the particles within some fixed volume in phase space. Let $V \subset \mathcal{R}^3$ be an arbitrary volume, let $\Omega \subset \mathcal{S}^2$ be an arbitrary solid angle, and consider the cartesian product $V \times \Omega$. This product is a set consisting of all pairs (\mathbf{r}, ω) such that $\mathbf{r} \in V$ and $\omega \in \Omega$, as depicted in Figure (3). Our task is to determine how the number of particles in $V \times \Omega$ changes with time. At any time t , the number of particles in this volume, $N(t)$, can be obtained by integrating phase space density over the 5-dimensional volume:

$$N(t) = \int_{\Omega} \int_V n(\mathbf{r}, \omega, t) d\mathbf{r} d\omega, \quad (5)$$

which results in a dimensionless quantity. We now consider the time derivative of $N(t)$. In general, the number of particles in a given volume can change with time due to time-dependent properties of the medium and sources, or due to transient effects as the system moves toward equilibrium. However, we need only consider the *stationary* or *steady state* distributions since the speed of light is effectively infinite compared to rates of macroscopic motion in habitable environments. Thus, we can assume that equilibrium is reached instantaneously with respect to the geometric configuration of the environment at any point in time. Each volume then has a fixed number of particles, which implies that phase space flux is independent of time, $\phi = \phi(\mathbf{r}, \omega)$, and that

$$\frac{dN(t)}{dt} = 0 \quad \left[\frac{1}{s} \right]. \quad (6)$$

Yet even in the stationary case time is evident in the constant motion of the particles, implying that the equilibrium is dynamic. The flow of particles into and out of $V \times \Omega$ must exactly balance. To examine this balance it is instructive to first identify all possible physical processes by which N could change, and then group the terms into gains and losses. By equating the gains with the losses we arrive at a stationary balance equation for $V \times \Omega$. Each of the three processes, emission, streaming, and collisions, can change the number of particles within $V \times \Omega$, so we require that

$$\begin{bmatrix} \text{change} \\ \text{due to} \\ \text{Emission} \end{bmatrix} + \begin{bmatrix} \text{change} \\ \text{due to} \\ \text{Streaming} \end{bmatrix} + \begin{bmatrix} \text{change} \\ \text{due to} \\ \text{Collisions} \end{bmatrix} = 0. \quad (7)$$

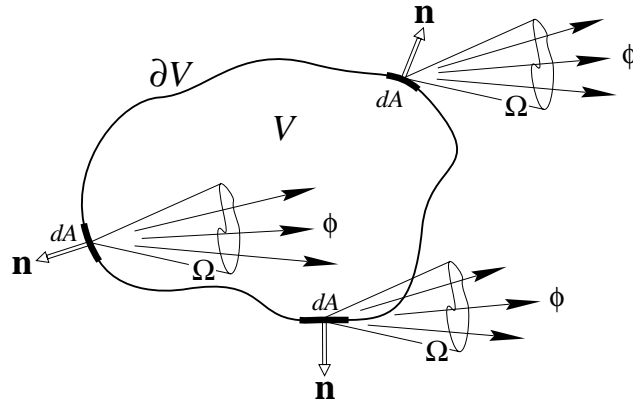


Figure 4: Integrating the flux due to Ω over the surface ∂V .

The first component, the change due to emission, accounts for the creation of new particles by sources. We assume that there are one or more physical processes generating new particles and that the combined effect is described by a *phase space source* function q , a real-valued function defined on $\mathcal{R}^3 \times \mathcal{S}^2$. The phase space source function gives the number of particles created per unit volume, per unit solid angle, per unit time, so its units are $\text{m}^{-3}\text{sr}^{-1}\text{s}^{-1}$. In terms of q , the change in N due to emission is

$$\mathbf{E} \equiv \int_{\Omega} \int_V q(\mathbf{r}, \omega) d\mathbf{r} d\omega \quad \left[\frac{1}{\text{s}} \right]. \quad (8)$$

Note that the units of \mathbf{E} are in accord with the time derivative of N . Next, we account for particles with directions in Ω that either escape from or enter into the volume V simply by streaming. More precisely, the change due to streaming is the net flow of particles with directions in Ω that pass through the surface ∂V of volume V . The process of streaming through a real or hypothetical surface is called *leakage*. The leakage through each differential patch on the surface depends only on the component of the flux that is normal to the patch. Integrating the normal component of the flux due to particles in Ω over the entire surface ∂V , we have

$$\mathbf{S} \equiv \int_{\Omega} \int_{\partial V} \phi(\mathbf{s}, \omega) \omega \cdot \mathbf{n}(\mathbf{s}) ds d\omega \quad \left[\frac{1}{\text{s}} \right], \quad (9)$$

which is the change in N due to streaming through ∂V . In general, some of the differential surface patches will have positive flows and some negative, as shown in Figure (4), so \mathbf{S} accounts for particles flowing both into and out of the volume V . With outward-pointing surface normals positive leakage indicates a net flow out of V , which decreases N .

The last component to analyze is the change due to collisions, which we separate into absorption and scattering. As discussed in section 2, we assume that particles interact solely with the medium and not each other, making the collision rate independent of ϕ . Analysis of the mechanisms of collisions belong to the realm of physical optics, quantum mechanics, etc. At the scale of particle transport, we simply incorporate models for the microscopic processes and look to other fields to supply the various constants for specific materials.

We assume that the probability of a particle disappearing due to absorption is proportional to the distance of travel through the medium. See Figure (5a). The constant of

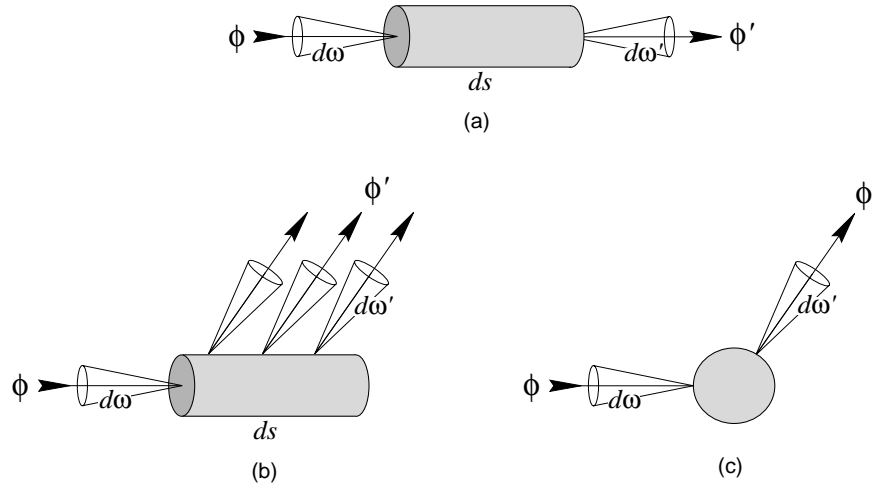


Figure 5: An intuitive view of volume interactions (a) per unit length, $\phi' \propto \phi ds$, (b) per unit solid angle per unit length, $\phi' \propto \phi d\omega' ds$, and (c) per unit solid angle, $\phi' \propto \phi d\omega'$.

proportionality is called the *absorption coefficient* σ_a , which has units of m^{-1} . The absorption coefficient may vary with position, but is normally assumed to be independent of direction. This is equivalent to assuming that the medium is isotropic, ruling out regular structure such as that of a crystal. The change in N due to absorption is expressed as

$$\mathbf{C}_{\text{abs}} \equiv \int_{\Omega} \int_V \sigma_a(\mathbf{r}) \phi(\mathbf{r}, \omega) d\mathbf{r} d\omega. \quad (10)$$

Collisions due to scattering are more complex. As with absorption, the probability of such a collision occurring is proportional to the distance the particle travels. Upon collision, however, the new direction must also be accounted for. Because of the one-speed assumption, we need not consider other changes to the particle as a result of scattering. We characterize the process of scattering by the *volume scattering kernel* k . The function k is defined on $\mathcal{R}^3 \times [-1, 1]$ such that

$$k(\mathbf{r}; \omega \cdot \omega') d\mathbf{r} d\omega'$$

is the probability that a particle at \mathbf{r} moving in the direction ω will be deflected into the new direction ω' , per unit solid angle about ω' per unit distance of travel. See Figure (5b). The units of k are $\text{m}^{-1}\text{sr}^{-1}$. That k depends only on the dot product $\omega \cdot \omega'$ and not the individual directions is again a statement about the isotropy of the medium.

Scattering may remove particles from $V \times \Omega$ by a process known as *out-scatter*. In out-scatter, a particle remains within V while its direction is instantaneously changed to something outside of Ω as a result of a collision. To account for out-scatter we define the term \mathbf{C}_{out} by

$$\mathbf{C}_{\text{out}} \equiv \int_{\Omega} \int_V \int_{S^2} k(\mathbf{r}; \omega \cdot \omega') \phi(\mathbf{r}, \omega) d\omega' d\mathbf{r} d\omega, \quad (11)$$

which is the total number of scattering collisions suffered by particles in $V \times \Omega$ per unit time. The opposite of out-scatter is a process by which particles are added to $V \times \Omega$ by

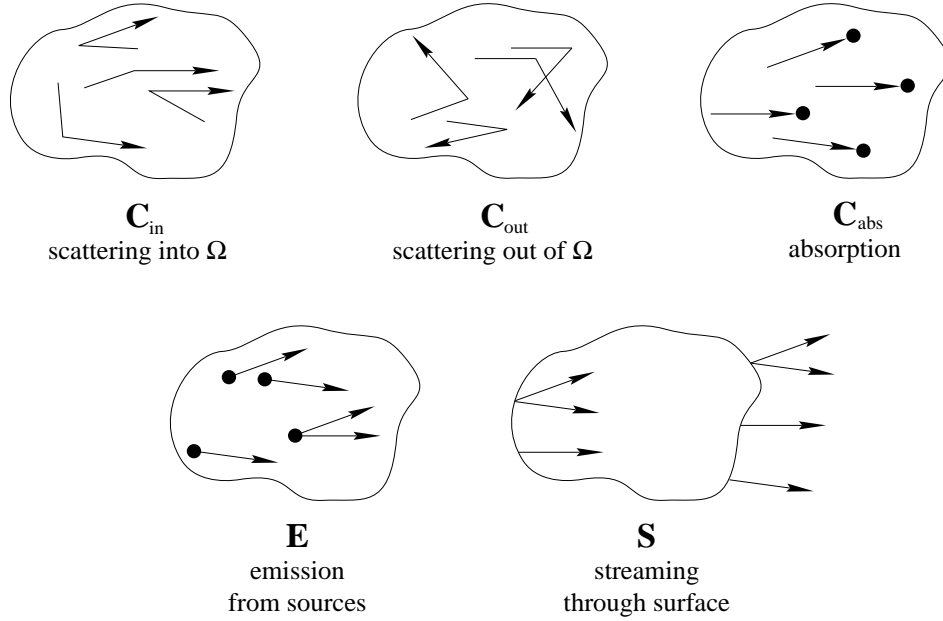


Figure 6: *The five processes affecting the number of particles in volume V with directions in solid angle Ω .*

being deflected into it. To account for *in-scatter* we define the term \mathbf{C}_{in} by

$$\mathbf{C}_{\text{in}} \equiv \int_{\Omega} \int_V \int_{\mathcal{S}^2} k(\mathbf{r}; \omega' \cdot \omega) \phi(\mathbf{r}, \omega') d\omega' d\mathbf{r} d\omega, \quad (12)$$

which is the number of particles in V , per unit time, with directions of travel in Ω after scattering. Note that both \mathbf{C}_{out} and \mathbf{C}_{in} account for particles whose directions lie within Ω both before and after scattering, which is neither a gain nor a loss. However, we need only consider the net change due to scattering, which is the difference between \mathbf{C}_{in} and \mathbf{C}_{out} , and this cancels the unwanted effect. Including the deflections from Ω back into Ω in both terms simplifies the definitions, allowing the integration to be over \mathcal{S}^2 instead of $\mathcal{S}^2 - \Omega$.

We can now collect the terms corresponding to the five distinct physical processes that can change N . Separating the terms into gains and losses and equating them, we have

$$\underbrace{\mathbf{S} + \mathbf{C}_{\text{abs}} + \mathbf{C}_{\text{out}}}_{\text{losses}} = \underbrace{\mathbf{E} + \mathbf{C}_{\text{in}}}_{\text{gains}} \quad (13)$$

The meanings of the five terms are summarized in Figure (6).

As it stands, equation (13) applies only to the phase space volume $V \times \Omega$, yet it contains within it an equation that applies globally. To extract the global balance equation we need only eliminate references to the fixed sets V and Ω . To do this, observe that all but one of the terms in equation (13) entail integration over V ; only the \mathbf{S} term is expressed as an integral over ∂V . Gauss's theorem can be used to convert the single surface integral into a volume integral, yielding

$$\mathbf{S} = \int_{\Omega} \int_V \omega \cdot \nabla \phi(\mathbf{r}, \omega) d\mathbf{r} d\omega. \quad (14)$$

Here, in addition to Gauss's theorem, we have used the fact that

$$\nabla \cdot [\omega \phi(\mathbf{r}, \omega)] = \omega \cdot \nabla \phi(\mathbf{r}, \omega)$$

which holds because ω is fixed with respect to the divergence operator. With the new expression for \mathbf{S} , all five terms in equation (13) entail integration over both V and Ω . Because $V \times \Omega$ is arbitrary, it follows that equality must hold for the integrands (to within a set of measure zero). Removing the two outer integrals from all five terms, we have

$$\begin{aligned} \omega \cdot \nabla \phi(\mathbf{r}, \omega) + \sigma_a(\mathbf{r}) \phi(\mathbf{r}, \omega) + \int_{\mathcal{S}^2} k(\mathbf{r}; \omega \cdot \omega') \phi(\mathbf{r}, \omega) d\omega' \\ = q(\mathbf{r}, \omega) + \int_{\mathcal{S}^2} k(\mathbf{r}; \omega' \cdot \omega) \phi(\mathbf{r}, \omega') d\omega'. \end{aligned} \quad (15)$$

Now, the integral corresponding to \mathbf{C}_{out} on the left-hand side of equation (15) can be simplified because the function ϕ does not depend on the variable of integration. We introduce the *scattering coefficient* σ_s defined by

$$\sigma_s(\mathbf{r}) \equiv \int_{\mathcal{S}^2} k(\mathbf{r}; \omega_0 \cdot \omega') d\omega' \quad \left[\frac{1}{\text{m}} \right],$$

which is the probability that a particle will suffer a scattering collision per unit distance traveled, in exact analogy with $\sigma_a(\mathbf{r})$. Here ω_0 can be an arbitrary direction in \mathcal{S}^2 because of isotropy. Then σ_s becomes a factor of ϕ in the balance equation. The two such factors, σ_s and σ_a , can be conveniently grouped into a *total interaction coefficient* σ ,

$$\sigma(\mathbf{r}) \equiv \sigma_s(\mathbf{r}) + \sigma_a(\mathbf{r}) \quad \left[\frac{1}{\text{m}} \right],$$

which is the probability that a particle will suffer either kind of collision, per unit distance traveled. With the new notation, equation (15) becomes

$$\boxed{\omega \cdot \nabla \phi(\mathbf{r}, \omega) + \sigma(\mathbf{r}) \phi(\mathbf{r}, \omega) = q(\mathbf{r}, \omega) + \int_{\mathcal{S}^2} k(\mathbf{r}; \omega' \cdot \omega) \phi(\mathbf{r}, \omega') d\omega'} \quad (16)$$

which is the standard form of the stationary one-speed particle transport equation in integro-differential form. When time-dependent phenomena are included, as in many nuclear and astrophysical applications, equation (16) appears in a slightly more general form; phase space flux becomes a function of time, $\phi = \phi(\mathbf{r}, \omega, t)$, and an additional term,

$$\frac{1}{v} \frac{\partial \phi(\mathbf{r}, \omega, t)}{\partial t}, \quad (17)$$

appears on the left-hand side, since equation (6) no longer holds.

The above derivation of the transport equation is one of many possible approaches. In the context of radiative transfer, Chandrasekhar [7] begins by considering directional derivatives of the flux, and Pomraning [40] describes a ‘‘Lagrangian’’ approach based on following packets of photons. Davison [8] and Weinberg and Wigner [59] elaborate many of these same ideas in the context of neutron transport theory. Case and Zweifel [6] and Kourganoff [23] give derivations that apply to many different physical phenomena. Finally, Preisendorfer [41, 42] has given a rigorous axiomatic derivation of the general transport equation based on measure theory.

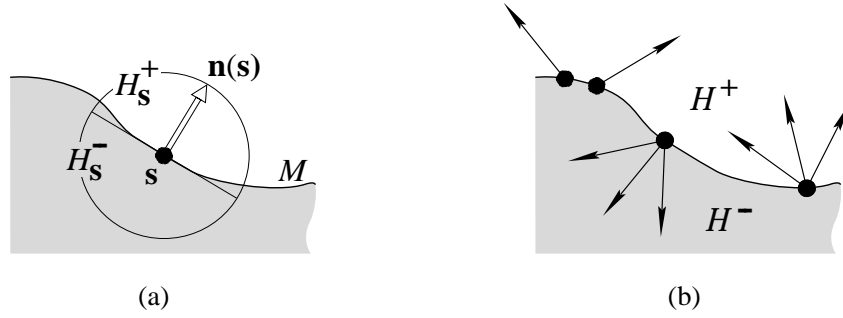


Figure 7: The surface M partitions (a) the set of directions \mathcal{S}^2 at each point $\mathbf{s} \in M$, and (b) the set of all rays originating on M .

2.2 Boundary Conditions

The presence of the gradient operator makes equation (16) a first-order differential equation in the spatial variables. As such, it cannot be a complete description of the physical situation in itself. A property of all differential equations is that they require boundary conditions to eliminate the arbitrary constants of integration [11].

Equation (16) is only valid away from surfaces, which are the boundaries. At the surfaces, collectively denoted M , we need to specify what happens. To express the boundary conditions, first observe that M imposes a natural partitioning of \mathcal{S}^2 into two hemispheres at each point. Let $\mathbf{n}(\mathbf{s})$ denote the surface normal at $\mathbf{s} \in M$. Then we designate the hemisphere where $\mathbf{n}(\mathbf{s}) \cdot \boldsymbol{\omega} > 0$ the *positive hemisphere* and define the set

$$H_{\mathbf{s}}^+ \equiv \{\boldsymbol{\omega} \in \mathcal{S}^2 : \boldsymbol{\omega} \cdot \mathbf{n}(\mathbf{s}) \geq 0\}.$$

Similarly, we define the collection of all pairs $(\mathbf{s}, \boldsymbol{\omega})$ corresponding to rays that originate on M and point outward:

$$\mathcal{H}^+ \equiv \{(\mathbf{s}, \boldsymbol{\omega}) \in M \times \mathcal{S}^2 : \boldsymbol{\omega} \in H_{\mathbf{s}}^+\}.$$

The sets $H_{\mathbf{s}}^-$ and \mathcal{H}^- are defined analogously, and satisfy

$$\begin{aligned} H_{\mathbf{s}}^+ \cup H_{\mathbf{s}}^- &= \mathcal{S}^2 \\ \mathcal{H}^+ \cup \mathcal{H}^- &= M \times \mathcal{S}^2. \end{aligned}$$

See Figure (7). The purpose of boundary conditions is to specify the behavior of the particles on the set \mathcal{H}^+ ; the set \mathcal{H}^- is accounted for by particles impinging on M . There are numerous ways in which these conditions can be specified, and different physical situations call for different boundary conditions [9, 25]. The simplest in form are the *explicit* boundary conditions in which

$$\phi(\mathbf{s}, \boldsymbol{\omega}) = q_b(\mathbf{s}, \boldsymbol{\omega}) \tag{18}$$

for all $(\mathbf{s}, \boldsymbol{\omega}) \in \mathcal{H}^+$. Explicit boundary conditions are independent of ϕ itself, accounting for particles that are generated by independent processes and subsequently enter the system

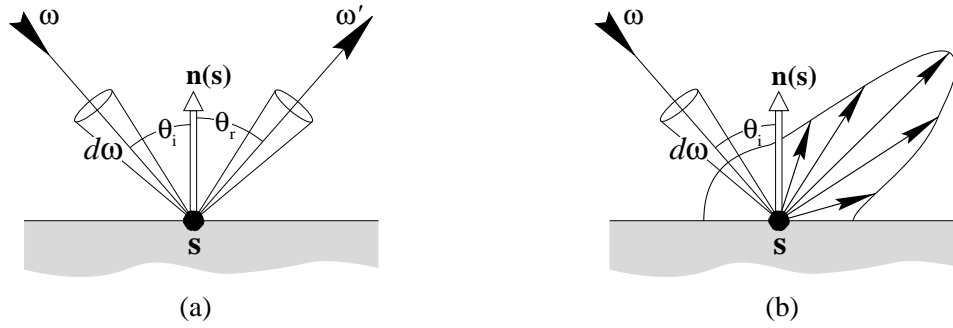


Figure 8: *Two types of implicit boundary conditions: (a) specular reflection, and (b) directional diffuse reflection.*

through the boundaries. In contrast, *implicit* or *reflecting* boundary conditions are of the form

$$\phi(\mathbf{s}, \omega) = f(\phi, \mathbf{s}, \omega) \quad (19)$$

for all $(\mathbf{s}, \omega) \in \mathcal{H}^+$. There are two common forms of implicit boundary conditions, both expressing ϕ restricted to \mathcal{H}^+ in terms ϕ restricted to \mathcal{H}^- . The physical interpretation is that of scattering at the boundary; particles originating at the boundary are a direct result of those impinging on it. The first form is *specular* reflection, and is of the form

$$\phi(\mathbf{s}, \omega) = F(\mathbf{s}, -\omega') \phi(\mathbf{s}, -\omega') \quad (20)$$

where ω' is the reflection of ω through the surface normal at \mathbf{s} . See Figure (8a). A second, and far more general type of implicit boundary condition is given by an integral transformation:

$$\phi(\mathbf{s}, \omega) = \int_{\mathcal{H}^-} k_b(\mathbf{s}; \omega' \rightarrow \omega) \phi(\mathbf{s}, \omega') d\omega', \quad (21)$$

where k_b is the *surface scattering kernel* with units of sr^{-1} . The arrow notation “ $\omega' \rightarrow \omega$ ” is used here in lieu of “ ω', ω ” to indicate incident and reflected directions. Note that the surface scattering kernel lacks the length dimension of volume scattering. This is because each particle scattering event takes place at the definite location on the surface determined by the direction of travel. The integral transformation of equation (21) expresses the density of particles reflected in any direction as a weighted “sum” of the incoming densities, and the weighting can depend strongly on both the incoming and outgoing directions. See Figure (8b). This phenomenon is also known as *directional diffuse* reflection [16]. Note that if we allow the kernel k_b to include *generalized functions*, such as the Dirac delta function, then we can write

$$k_b(\mathbf{s}; \omega' \rightarrow \omega) = F(\mathbf{s}, \omega') \delta(\omega' - \omega),$$

which subsumes the case of specular reflection. Finally, we note that there are physical constraints on both F and k_b . If the only source of new particles is q , conservation dictates the constraints

$$F(\mathbf{s}, \omega) \leq 1 \quad (22)$$

and

$$\int_{H_{\mathbf{s}}^-} k_b(\mathbf{s}; \omega' \rightarrow \omega) d\omega(\omega') \leq 1 \quad (23)$$

for all $(\mathbf{s}, \omega) \in \mathcal{H}^+$. Similarly, both F and k_b must be non-negative everywhere.

3 The Equation of Transfer

The general transport equation derived in the last section is not tied to a particular physical situation. We now rephrase it slightly to apply to the transfer of radiant energy; that is, we now account for the semantics of photons. There are three properties of photons that need to be incorporated into the equation. First, photons travel at a constant speed; in a vacuum, this speed is a universal constant denoted by c . Second, photons have an associated frequency, an important internal attribute that determines their energy and affects their behavior. And finally, trajectories of photons are strongly influenced by surfaces; a property shared by gas molecules, but not neutrons.

We begin by translating the abstract notion of phase space flux into the corresponding radiometric concept of *radiance*. Radiance is the power per unit area per unit solid angle due to radiant energy crossing a surface perpendicularly (or, equivalently, per unit *projected* area). This is essentially the same concept as phase space flux; all that is missing is the notion of energy per photon. But the energy of a photon is related to its frequency by

$$E = h\nu \quad [\text{J}], \quad (24)$$

where h is Planck's constant. It follows that the relationship between phase space flux and radiance $L(\mathbf{r}, \omega)$ is

$$\begin{aligned} L(\mathbf{r}, \omega) &= h\nu \phi(\mathbf{r}, \omega) \\ &= ch\nu n(\mathbf{r}, \omega) \quad \left[\frac{\text{W}}{\text{m}^2 \text{ sr}} \right]. \end{aligned}$$

Similarly, we need to translate the phase space source function into a quantity related to radiant power. We define a volume source term ε by

$$\varepsilon(\mathbf{r}, \omega) \equiv h\nu q(\mathbf{r}, \omega) \quad \left[\frac{\text{W}}{\text{m}^3 \text{ sr}} \right].$$

The physical process of emission is a phenomenon of matter, and therefore applies only to volumes. Nonetheless, it is a useful abstraction to allow surfaces to be emitters [39]. We define a surface source term ε_b by

$$\varepsilon_b(\mathbf{s}, \omega) \equiv h\nu q_b(\mathbf{s}, \omega) \quad \left[\frac{\text{W}}{\text{m}^2 \text{ sr}} \right].$$

The functions L , ε , and ε_b will play the same roles as their abstract counterparts, with the difference that they are now linked to physical units meaningful to the transfer of radiant energy.

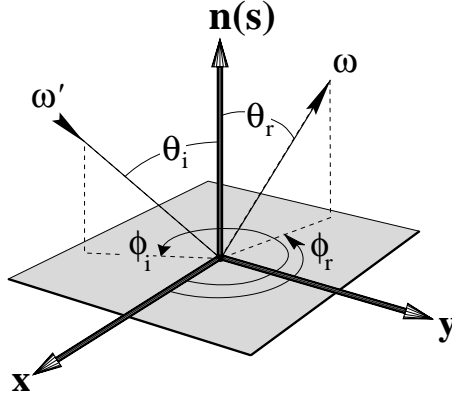


Figure 9: *The local coordinates for bidirectional reflectance.*

Because surfaces are a dominant influence in the realm of photons, we impose boundary conditions that are most general, including both explicit and implicit terms. Then the transport equation can be written in terms of L :

$$\begin{aligned}
 \omega \cdot \nabla L(\mathbf{r}, \omega) + \sigma(\mathbf{r})L(\mathbf{r}, \omega) &= \varepsilon(\mathbf{r}, \omega) + \int_{S^2} k(\mathbf{r}; \omega' \cdot \omega) L(\mathbf{r}, \omega') d\omega' \\
 L(\mathbf{s}, \omega) &= \varepsilon_b(\mathbf{s}, \omega) + \int_{H_{\mathbf{s}}^-} k_b(\mathbf{s}; \omega' \rightarrow \omega) L(\mathbf{s}, \omega') d\omega'
 \end{aligned} \tag{25}$$

where \mathbf{s} is a boundary point. Equation (25) is the stationary gray equation of transfer with boundary conditions.

Both emission and collision behaviors of neutral particles may depend upon internal states or speed. Therefore $\varepsilon, \varepsilon_b, k$, and k_b are generally all functions of these attributes. In photon transport these functions are usually strongly frequency-dependent. The one-speed assumption, which translates into fixed frequency for photons, allows simulations at different frequencies to be decoupled. It is common in global illumination to use three or more concurrent but independent simulations corresponding to different frequencies.

The surface scattering kernel may be expressed in radiometric terms by relating it to the bidirectional reflectance distribution function, or BRDF, ρ . Denoting the BRDF at the point $\mathbf{s} \in M$ by $\rho_{\mathbf{s}}$, we have

$$k_b(\mathbf{s}, \omega' \rightarrow \omega) = \rho_{\mathbf{s}}(\theta_i, \phi_i, \theta_r, \phi_r) \cos \theta_i \quad \left[\frac{1}{\text{sr}} \right], \tag{26}$$

where the incoming and outgoing directions have been expressed as incident and reflected angles to indicate a shift to local coordinates. For instance, $\cos \theta_i = -\mathbf{n}(\mathbf{s}) \cdot \omega'$ is the cosine of the angle of incidence. See Figure (9). The units of $\rho_{\mathbf{s}}$ are also sr^{-1} , as $\cos \theta_i$ is dimensionless. The volume scattering kernel can also be put into another form by factoring out the *phase function* $P_{\mathbf{r}}(x)$ defined by

$$P_{\mathbf{r}}(x) \equiv 4\pi \frac{k(\mathbf{r}, x)}{\sigma_s(\mathbf{r})} \quad \left[\frac{1}{\text{sr}} \right], \tag{27}$$

where $x \in [-1, 1]$ encodes the cosine of the angle between the directions before and after scattering. In terms of $P_{\mathbf{r}}$, the volume scattering kernel is

$$k(\mathbf{r}, \omega' \cdot \omega) = \sigma_s(\mathbf{r}) \frac{P_{\mathbf{r}}(\omega' \cdot \omega)}{4\pi}, \quad (28)$$

which has the advantage of separating the notions of collision probability per unit length and the resulting distribution of directions. See Figure (5c). The 4π appears in the definition of phase function so that $P \equiv 1$ corresponds to uniform scattering in all directions. A common empirical phase function due to Rayleigh is

$$P(x) = \frac{3}{4}(1 + x^2). \quad (29)$$

3.1 Integral Form

In this section we derive an alternate form of the equation of transfer that is a pure integral equation instead of an integro-differential equation, as equation (25). In performing this transformation the boundary conditions are absorbed into the equation, leaving a single expression that is equivalent but sometimes more convenient. The derivation below parallels those of Pomraning [40] and Williams [63]. The idea is to integrate equation (25) along a ray until a boundary point is reached. The integration eliminates the differential operator and also incorporates the boundary value at the point of intersection with the ray. The key observation is that the operator $\omega \cdot \nabla$ is a directional derivative in the direction ω , allowing us to write

$$\begin{aligned} \omega \cdot \nabla L(\mathbf{r}, \omega) &= \left. \frac{\partial}{\partial x} L(\mathbf{r} + x\omega, \omega) \right|_{x=0} \\ &= - \left. \frac{\partial}{\partial x} L(\mathbf{r} - x\omega, \omega) \right|_{x=0} \end{aligned}$$

where $x \in \mathcal{R}$ parametrizes the ray through \mathbf{r} parallel to ω . We shall express $L(\mathbf{r}, \omega)$ in terms of what happens along the ray $(\mathbf{r}, -\omega)$, which makes the second expression above a more convenient choice. To simplify notation, we define $Q(\mathbf{r}, \omega)$ to be the gains due to volume emission and in-scattering,

$$Q(\mathbf{r}, \omega) \equiv \varepsilon(\mathbf{r}, \omega) + \int_{S^2} k(\mathbf{r}; \omega' \cdot \omega) L(\mathbf{r}, \omega') d\omega', \quad (30)$$

and introduce the single-parameter functions

$$\begin{aligned} \hat{L}(x) &\equiv L(\mathbf{r} - x\omega, \omega) \\ \hat{Q}(x) &\equiv Q(\mathbf{r} - x\omega, \omega) \\ \hat{\sigma}(x) &\equiv \sigma(\mathbf{r} - x\omega), \end{aligned}$$

where \mathbf{r} and ω are to be considered fixed. Ignoring the boundary conditions for the moment, equation (25) can be written as

$$\frac{d}{dx} \hat{L}(x) - \hat{\sigma}(x) \hat{L}(x) = -\hat{Q}(x). \quad (31)$$

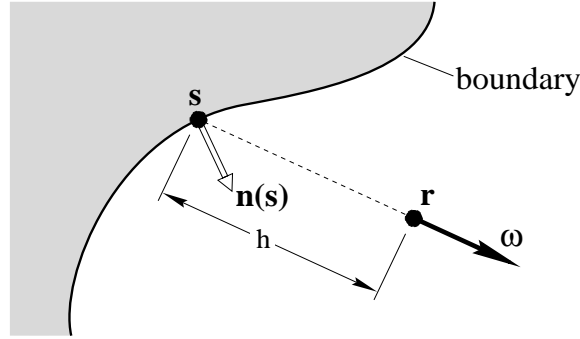


Figure 10: The boundary point s corresponding to the ray (\mathbf{r}, ω) . The point s is given by $\gamma(\mathbf{r}, -\omega)$, and the distance h by $b(\mathbf{r}, -\omega)$.

Note that the term $\hat{Q}(x)$ has no explicit dependence on \hat{L} . Even though \hat{L} is subsumed by the in-scattering term of equation (30), its contribution is only on a set of measure zero. Therefore, equation (31) is a linear first-order ordinary differential equation in \hat{L} with variable coefficients, which is easily solved by means of an integrating factor [45]. By defining

$$v(x) \equiv \exp \left[- \int_0^x \hat{\sigma}(y) dy \right], \quad (32)$$

equation (31) can be written as

$$\frac{d}{dx} [v(x) \hat{L}(x)] = -v(x) \hat{Q}(x), \quad (33)$$

which is in a form that can be integrated directly. By integrating both sides from 0 to x and noting that $v(0) = 1$, we solve for $\hat{L}(0)$, yielding

$$\hat{L}(0) = v(x) \hat{L}(x) + \int_0^x v(y) \hat{Q}(y) dy. \quad (34)$$

Here we have chosen to solve for $\hat{L}(0)$ because $\hat{L}(0) = L(\mathbf{r}, \omega)$. Although equation (34) holds for any x , it is only useful if we have a convenient expression for $\hat{L}(x)$; otherwise we have made no progress toward a simpler representation for $\hat{L}(0)$. Fortunately, if we choose x so that $\hat{L}(x)$ lies on the boundary, the boundary conditions provide such an expression. That is, we let $x = b(\mathbf{r}, -\omega)$ where

$$b(\mathbf{r}, \omega) \equiv \inf \{x > 0 : \mathbf{r} + x\omega \in M\}$$

gives the distance to the nearest point of intersection with the boundary along the given ray. Similarly, we define the closely related function γ by

$$\gamma(\mathbf{r}, \omega) \equiv \mathbf{r} + b(\mathbf{r}, \omega)\omega,$$

which returns the actual point of intersection. See Figure (10). For simplicity, we assume that the surfaces M form a complete enclosure so that b and γ are always well-defined; this

is always possible by surrounding the scene with a non-reflecting surface if necessary. Now, to incorporate the value at the boundary into equation (34), we define Γ to be the boundary function

$$\Gamma(\mathbf{s}, \omega) \equiv \varepsilon_b(\mathbf{s}, \omega) + \int_{H_{\mathbf{s}}^-} k_b(\mathbf{s}; \omega' \rightarrow \omega) L(\mathbf{s}, \omega') d\omega' \quad (35)$$

for any $(\mathbf{s}, \omega) \in \mathcal{H}^+$. We must also generalize equation (34) back to a form that holds for any ray, so we introduce several new functions. The *optical distance* function is defined by

$$\tau(\mathbf{r}, \mathbf{r}') \equiv \int_0^{|\mathbf{r}' - \mathbf{r}|} \sigma(\mathbf{r} + y\mathbf{u}) dy, \quad (36)$$

where $\mathbf{u} = (\mathbf{r}' - \mathbf{r}) / \|\mathbf{r}' - \mathbf{r}\|$ is the unit vector pointing from \mathbf{r} to \mathbf{r}' , and the *path absorption* function is defined by

$$\beta(\mathbf{r}, \mathbf{r}') \equiv e^{-\tau(\mathbf{r}, \mathbf{r}')}. \quad (37)$$

The function β is simply a more general form of the integrating factor $v(x)$. Note that both τ and β are dimensionless. Finally, the stationary gray equation of transfer in integral form can be written as

$$L(\mathbf{r}, \omega) = \beta(\mathbf{s}, \mathbf{r}) \Gamma(\mathbf{s}, \omega) + \int_0^h \beta(\mathbf{r} - x\omega, \mathbf{r}) Q(\mathbf{r} - x\omega, \omega) dx \quad (38)$$

where $\mathbf{s} = \gamma(\mathbf{r}, -\omega)$ and $h = b(\mathbf{r}, -\omega)$. Note that both Γ and Q are functions of L , so equation (38) is not a closed-form expression for L ; it is merely an alternative to equation (25) in which the boundary conditions are built in. Similar equations can be derived when L is time-dependent [63] and where there is coupling between frequencies [58]; in both cases the derivation proceeds as above after applying a Laplace transform.

3.2 Special Cases of the Equation of Transfer

The equation of transfer has several direct applications in global illumination, including the simulating of participating media [21, 46] and physically similar effects [24] as well as subsurface scattering [15]. But its usefulness extends beyond these applications. Because of the generality of the equation of transfer, it subsumes many of the equations used in global illumination and volume rendering. Just as the rendering equation [19] unifies an array of rendering techniques, the equation of transfer encompasses an even larger class. We now consider some of the special cases that it subsumes.

3.2.1 Vacuum Conditions

If the space separating the surfaces is a vacuum, there can be no volume emission and no particle collisions except at surfaces. Under these conditions equation (25) all but disappears, leaving behind little more than the boundary conditions. More precisely, vacuum conditions imply that

$$\begin{aligned} \varepsilon &\equiv 0 \\ \sigma_a &\equiv 0 \\ \sigma_s &\equiv 0, \end{aligned}$$

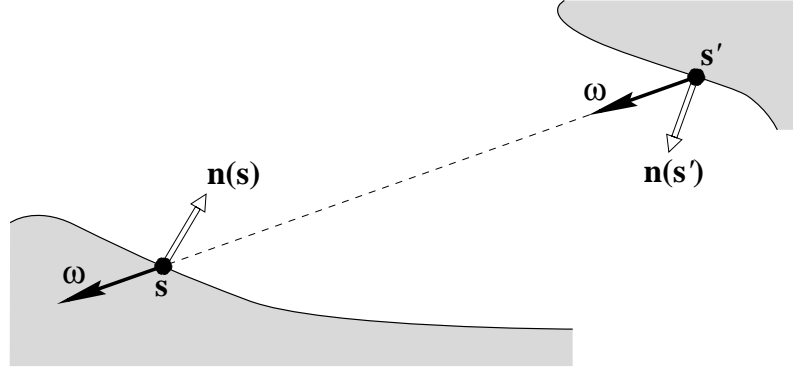


Figure 11: *With vacuum conditions, the incident radiance at (\mathbf{s}, ω) is equal to the radiance at (\mathbf{s}', ω) .*

therefore, equation (25) reduces to

$$\omega \cdot \nabla L(\mathbf{r}, \omega) = 0 \quad (39)$$

for all $\mathbf{r} \in \mathcal{R}^3 - M$ and $\omega \in \mathcal{S}^2$. Equation (39) states that the directional derivative of L vanishes everywhere in a vacuum. This implies that the radiance function $L(\mathbf{r}, \omega)$ is invariant along each direction ω , and consequently, $L(\mathbf{r}, \omega)$ at any interior point is determined by a point on the boundary:

$$L(\mathbf{r}, \omega) = L(\gamma(\mathbf{r}, -\omega), \omega) \quad (40)$$

for all $(\mathbf{r}, \omega) \in \mathcal{R}^3 \times \mathcal{S}^2$. See Figure (11). Substituting equation (40) into the boundary condition of equation (25), we have

$$L(\mathbf{s}, \omega) = \varepsilon_b(\mathbf{s}, \omega) + \int_{H_{\mathbf{s}}^-} k_b(\mathbf{s}; \omega' \rightarrow \omega) L(\gamma(\mathbf{s}, -\omega'), \omega') d\omega' \quad (41)$$

which is a balance equation for $L(\mathbf{s}, \omega)$ restricted to \mathcal{H}^+ . This is the governing equation for direct radiative transfer between surfaces; it describes the same physical situation as the rendering equation [19], but in terms of radiance.

The equation simplifies further if we assume that all surfaces are pure diffuse reflectors. In this case the radiance function depends on position but not direction. Making the corresponding changes to L and ρ , we arrive at the continuous version of the radiosity equation:

$$L(\mathbf{s}) = \varepsilon_b(\mathbf{s}) + \rho(\mathbf{s}) \int_{H_{\mathbf{s}}^-} L(\gamma(\mathbf{s}, -\omega')) \cos \theta_i d\omega'. \quad (42)$$

3.2.2 Non-Scattering Media

Now consider the case where there is a medium that absorbs and emits, but does not scatter. That is,

$$\sigma_s \equiv 0.$$

In this case it will be easier to study the implication by starting with the integral form of the equation of transfer. The absence of scattering implies that the function Q defined in equation (30) reduces to

$$Q(\mathbf{r}, \omega) \equiv \varepsilon(\mathbf{r}, \omega) \quad (43)$$

so the governing equation is

$$L(\mathbf{r}, \omega) = \beta(\mathbf{s}, \mathbf{r}) \Gamma(\mathbf{s}, \omega) + \int_0^h \beta(\mathbf{r} - x\omega, \mathbf{r}) \varepsilon(\mathbf{r} - x\omega, \omega) dx \quad (44)$$

where $\mathbf{s} = \gamma(\mathbf{r}, -\omega)$ and $h = b(\mathbf{r}, -\omega)$. If we further assume that the boundaries are non-reflecting, then

$$\Gamma(\mathbf{s}, \omega) \equiv \varepsilon_b(\mathbf{s}, \omega) \quad (45)$$

for all $(\mathbf{s}, \omega) \in \mathcal{H}^+$. Under these assumptions equation (44) no longer depends upon L at all. The value of $L(\mathbf{r}, \omega)$ is therefore a function only of the absorption coefficient $\sigma_a(\mathbf{r})$ and the volume emittance function $\varepsilon(\mathbf{r}, \omega)$ along the path of the ray $(\mathbf{r}, -\omega)$. Using the notation introduced in section 3.1, the equation for $L(\mathbf{r}, \omega)$ can then be written as

$$\begin{aligned} \hat{L}(0) &= v(x) \hat{\varepsilon}_b(x) + \int_0^x v(y) \hat{\varepsilon}(y) dy \\ &= \exp \left[- \int_0^x \hat{\sigma}(z) dz \right] \hat{\varepsilon}_b(x) + \int_0^x \exp \left[- \int_0^y \hat{\sigma}(z) dz \right] \hat{\varepsilon}(y) dy \end{aligned} \quad (46)$$

where x is the distance to the first surface, and

$$\begin{aligned} \hat{\varepsilon}_b(x) &\equiv \varepsilon_b(\mathbf{r} - x\omega, \omega) \\ \hat{\varepsilon}(x) &\equiv \varepsilon(\mathbf{r} - x\omega, \omega). \end{aligned}$$

Finally, if we neglect the surfaces altogether, we have the emission-absorption model frequently used in volume rendering [21]:

$$\hat{L}(0) = \int_0^\infty \exp \left[- \int_0^y \hat{\sigma}(z) dz \right] \hat{\varepsilon}(y) dy. \quad (47)$$

Nomenclature

<i>Symbol</i>	<i>Units</i>	<i>Page</i>	<i>Meaning</i>
$b(\mathbf{r}, \omega)$	m	1-17	distance along the ray to boundary
$H_{\mathbf{s}}^+$	–	1-12	positive hemisphere at a point $\mathbf{s} \in M$
\mathcal{H}^+	–	1-12	outgoing rays originating on M
$k(\mathbf{r}; \omega \cdot \omega')$	$\text{m}^{-1} \text{sr}^{-1}$	1-9	volume scattering kernel
$k_b(\mathbf{s}; \omega \rightarrow \omega')$	sr^{-1}	1-13	surface scattering kernel
$L(\mathbf{r}, \omega)$	$\text{W m}^{-2} \text{sr}^{-1}$	1-14	radiance function
M	–	1-12	all surfaces (boundaries)
$n(\mathbf{r}, \omega)$	$\text{m}^{-3} \text{sr}^{-1}$	1-5	phase space density
$\mathbf{n}(\mathbf{s})$	–	1-12	surface normal function
$P_{\mathbf{r}}(x)$	sr^{-1}	1-15	normalized phase function at \mathbf{r}
$q(\mathbf{r}, \omega)$	$\text{m}^{-3} \text{sr}^{-1} \text{s}^{-1}$	1-8	phase space source function
$q_b(\mathbf{s}, \omega)$	$\text{m}^{-2} \text{sr}^{-1} \text{s}^{-1}$	1-12	explicit boundary condition
\mathbf{r}	–	1-5	a point in \mathcal{R}^3 representing a position
\mathcal{R}^n	–	1-4	n -dimensional Euclidean space
\mathbf{s}	–	1-12	a point in \mathcal{R}^3 representing a boundary point
\mathcal{S}^2	–	1-4	the unit sphere in \mathcal{R}^3
V	–	1-7	a volume in \mathcal{R}^3
$\beta(\mathbf{r}, \mathbf{r}')$	–	1-18	path absorption
$\sigma_a(\mathbf{r})$	m^{-1}	1-9	absorption coefficient
$\sigma_s(\mathbf{r})$	m^{-1}	1-11	scattering coefficient
$\sigma(\mathbf{r})$	m^{-1}	1-11	total interaction coefficient
$\varepsilon(\mathbf{r}, \omega)$	$\text{W m}^{-3} \text{sr}^{-1}$	1-14	volume emittance function
$\varepsilon_b(\mathbf{s}, \omega)$	$\text{W m}^{-2} \text{sr}^{-1}$	1-14	surface emittance function
$\phi(\mathbf{r}, \omega)$	$\text{m}^{-2} \text{sr}^{-1} \text{s}^{-1}$	1-5	phase space flux
$\gamma(\mathbf{r}, \omega)$	–	1-17	ray shooting function
$\rho_{\mathbf{s}}(\theta_i, \phi_i, \theta_r, \phi_r)$	sr^{-1}	1-15	bidirectional reflectance distribution function
$\tau(\mathbf{r}, \mathbf{r}')$	–	1-18	optical distance
ω	–	1-5	a vector in \mathcal{S}^2 representing a direction
Ω	–	1-7	a subset of \mathcal{S}^2 representing a solid angle

References

- [1] James Arvo and David Kirk. Particle transport and image synthesis. *Computer Graphics*, 24(4):63–66, August 1990.
- [2] R. Byron Bird, Warren E. Stewart, and Edwin N. Lightfoot. *Transport Phenomena*. John Wiley & Sons, New York, 1960.
- [3] Christoph C. Borel, Siegfried A. W. Gerstl, and Bill J. Powers. The radiosity method in optical remote sensing of structured 3-d surfaces. *Remote Sensing of the Environment*, 36:13–44, 1991.
- [4] Chris Buckalew and Donald Fussell. Illumination networks: Fast realistic rendering with general reflectance functions. *Computer Graphics*, 23(3):89–98, July 1989.
- [5] Kenneth M. Case, F. de Hoffmann, and G. Placzek. *Introduction to the Theory of Neutron Diffusion*, volume 1. Los Alamos Scientific Laboratory, Los Alamos, New Mexico, 1953.
- [6] Kenneth M. Case and Paul F. Zweifel. *Linear Transport Theory*. Addison-Wesley, New York, 1967.

- [7] S. Chandrasekar. *Radiative Transfer*. Dover Publications, New York, 1960.
- [8] B. Davison. *Neutron Transport Theory*. Oxford University Press, New York, 1957.
- [9] James J. Duderstadt and William R. Martin. *Transport Theory*. John Wiley & Sons, New York, 1979.
- [10] Richard P. Feynmann. *QED, The Strange Theory of Light and Matter*. Princeton University Press, Princeton, New Jersey, 1985.
- [11] Samuel Glasstone and Alexander Sesonske. *Nuclear Reactor Engineering*. Van Nostrand, Princeton, New Jersey, 1967.
- [12] Narendra S. Goel, Ivan Rozehnal, and Richard L. Thompson. A computer graphics based model for scattering from objects of arbitrary shapes in the optical region. *Remote Sensing of the Environment*, 36:73–104, 1991.
- [13] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennett Battaile. Modeling the interaction of light between diffuse surfaces. *Computer Graphics*, 18(3):213–222, July 1984.
- [14] Harold Grad. Principles of the kinetic theory of gases. In S. Flügge, editor, *Handbuch der Physik*, volume 12, pages 205–294. Springer-Verlag, New York, 1958.
- [15] Xiao D. He. *Physically-Based Models for the Reflection, Transmission and Subsurface Scattering of Light by Smooth and Rough Surfaces, with Applications to Realistic Image Synthesis*. PhD thesis, Cornell University, January 1993.
- [16] Xiao D. He, Kenneth E. Torrance, François X. Sillion, and Donald P. Greenberg. A comprehensive physical model for light reflection. *Computer Graphics*, 25(4):175–186, July 1991.
- [17] W. L. Hendry, K. D. Lathrop, S. Vandervoort, and J. Wooten. Bibliography on neutral particle transport theory. Technical Report LA-4287-MS, United States Atomic Energy Commission, 1970.
- [18] J. C. Hunsaker. Dimensional analysis and similitude in mechanics. In *Theodore von Kármán Anniversary Volume; Contributions to Applied Mechanics and Related Subjects*, pages 21–48. California Institute of Technology, Pasadena, California, 1941.
- [19] James T. Kajiya. The rendering equation. *Computer Graphics*, 20(4):143–150, August 1986.
- [20] James T. Kajiya. Radiometry and photometry for computer graphics. In *Advanced Topics in Ray Tracing, SIGGRAPH 90 Course Notes*, volume 24, August 1990.
- [21] James T. Kajiya and Brian P. Von Herzen. Ray tracing volume densities. *Computer Graphics*, 18(3):165–173, July 1984.
- [22] Morris Kline and Irvin W. Kay. *Electromagnetic Theory and Geometrical Optics*. John Wiley & Sons, New York, 1965.
- [23] V. Kourganoff. *Basic Methods in Transfer Problems: Radiative Equilibrium and Neutron Diffusion*. Oxford University Press, New York, 1952.
- [24] Wolfgang Krueger. Volume rendering and data feature enhancement. *Computer Graphics*, 24(5):21–26, November 1990.
- [25] Jacqueline Lenoble, editor. *Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures*. Studies in Geophysical Optics and Remote Sensing. A. Deepak, Hampton, Virginia, 1985.
- [26] Jeffery Lewins. *Importance, The Adjoint Function: The Physical Basis of Variational and Perturbation Theory in Transport and Diffusion Problems*. Pergamon Press, New York, 1965.
- [27] E. E. Lewis and W. F. Miller, Jr. *Computational Methods of Neutron Transport*. John Wiley & Sons, New York, 1984.

- [28] Rudolf K. Luneburg. *Mathematical Theory of Optics*. University of California, Berkeley, California, 1964.
- [29] Donald H. Menzel, editor. *Selected Papers on the Transfer of Radiation*. Dover Publications, New York, 1966.
- [30] Jurgen R. Meyer-Arendt. Radiometry and photometry: Units and conversion factors. *Applied Optics*, 7(10):2081–2084, October 1968.
- [31] Parry Moon. *The Scientific Basis of Illuminating Engineering*. McGraw-Hill, New York, 1936.
- [32] Parry Moon and Domina E. Spencer. *Lighting Design*. Addison-Wesley, Cambridge, Massachusetts, 1948.
- [33] Ranga B. Myneni and Juhan Ross, editors. *Photon-Vegetation Interactions*. Springer-Verlag, New York, 1991.
- [34] Ranga B. Myneni, Juhan Ross, and Ghassem Asrar. A review on the theory of photon transport in leaf canopies. *Agricultural and Forest Meteorology*, 45(1-2):1–153, February 1989.
- [35] National Bureau of Standards. *The International System of Units (SI)*. U.S. Government Printing Office, Washington DC, 1977. Special Publication 330.
- [36] Fred E. Nicodemus. Radiance. *American Journal of Physics*, 31:368–377, 1963.
- [37] Fred E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. *Geometrical Considerations and Nomenclature for Reflectance*. Monograph number 160. National Bureau of Standards, Washington DC, 1977.
- [38] Tomoyuki Nishita and Eihachiro Nakamae. Continuous tone representation of 3-D objects taking account of shadows and interreflection. *Computer Graphics*, 19(3):23–30, July 1985.
- [39] Max Planck. *The Theory of Heat Radiation*. Dover Publications, New York, 1988.
- [40] G. C. Pomraning. *The Equations of Radiation Hydrodynamics*. Pergamon Press, New York, 1973.
- [41] Rudolph W. Preisendorfer. A mathematical foundation for radiative transfer theory. *Journal of Mathematics and Mechanics*, 6(6):685–730, November 1957.
- [42] Rudolph W. Preisendorfer. Radiative transfer axioms. Technical Report 57-44, Scripps Institution of Oceanography Visibility Laboratory, University of California, San Diego, October 1957.
- [43] Rudolph W. Preisendorfer. *Radiative Transfer on Discrete Spaces*. Pergamon Press, New York, 1965.
- [44] Rudolph W. Preisendorfer. *Hydrologic Optics, Volumes I–VI*. National Oceanic & Atmospheric Administration, Honolulu, Hawaii, 1976. (Available as NTIS PB-259 793).
- [45] Earl D. Rainville and Phillip E. Bedient. *Elementary Differential Equations*. The Macmillan Company, New York, fifth edition, 1974.
- [46] Holly E. Rushmeier and Kenneth E. Torrance. The zonal method for calculating light intensities in the presence of a participating medium. *Computer Graphics*, 21(4):293–302, July 1987.
- [47] Holly E. Rushmeier and Kenneth E. Torrance. Extending the radiosity method to include specularly reflecting and translucent materials. *ACM Transactions on Graphics*, 9(1):1–27, January 1990.
- [48] Arthur Schuster. Radiation through a foggy atmosphere. *Astrophysical Journal*, 21(1):1–22, January 1905.

- [49] K. Schwarzschild. On the equilibrium of the sun's atmosphere. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Math.-phys. Klasse*, 195:41–53, 1906. (See [29] for English translation).
- [50] Peter Shirley. A ray tracing algorithm for global illumination. *Graphics Interface '90*, pages 205–212, May 1990.
- [51] Robert Siegel and John R. Howell. *Thermal Radiation Heat Transfer*. Hemisphere Publishing Corp., New York, second edition, 1981.
- [52] François Sillion, James Arvo, Stephen Westin, and Donald P. Greenberg. A global illumination solution for general reflectance distributions. *Computer Graphics*, 25(4):187–196, July 1991.
- [53] François Sillion and Claude Puech. A general two-pass method integrating specular and diffuse reflection. *Computer Graphics*, 23(4), August 1989.
- [54] Brian Smits, James Arvo, and David Salesin. An importance-driven radiosity algorithm. *Computer Graphics*, 26(4):273–282, July 1992.
- [55] Jerome Spanier and Ely M. Gelbard. *Monte Carlo Principles and Neutron Transport Problems*. Addison-Wesley, Reading, Massachusetts, 1969.
- [56] I. F. Spiro, R. C. Jones, and D. Q. Wark. Atmospheric transmission: Concepts, symbols, units and nomenclature. *Infrared Physics*, 5(1):11–36, 1965.
- [57] Gregory J. Ward. Measuring and modeling anisotropic reflection. *Computer Graphics*, 26(2):265–272, July 1992.
- [58] Alvin M. Weinberg. *Theory of Neutron Chain Reactions*. Technical Information Service, Oak Ridge, Tennessee, 1951.
- [59] Alvin M. Weinberg and Eugene P. Wigner. *The Physical Theory of Neutron Chain Reactors*. University of Chicago Press, 1958.
- [60] Stephen Westin, James Arvo, and Kenneth Torrance. Predicting reflectance functions from complex surfaces. *Computer Graphics*, 26(2):255–264, July 1992.
- [61] Turner Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 32(6):343–349, June 1980.
- [62] Eugene P. Wigner. Mathematical problems of nuclear reactor theory. In *Nuclear Reactor Theory. Proceedings of the Eleventh Symposium in Applied Mathematics*, pages 89–104. American Mathematical Society, Providence, Rhode Island, 1961.
- [63] M. M. R. Williams. *Mathematical Methods in Particle Transport Theory*. John Wiley & Sons, New York, 1971.
- [64] Lawrence B. Wolf. Diffuse reflection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 23(4):472–478, June 1992.

Global Illumination:
Transfer Equations, Linear Operators and
Integral Equations

Jim Arvo
Program of Computer Graphics
Cornell University

The Equations of Global Illumination

- **Origins**
- **Assumptions**
- **Connections**


What we'll cover

- Equations of Radiative Transfer
 - Derivation of classical transfer equation*
 - Physical assumptions*
 - Connections with other fields*
 - Simplified models*
- Mathematical Tools & Concepts
 - Linear Operators*
 - Integral Equations*
 - Using "Importance"*

Physical Models and
Physically-Based Rendering

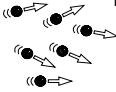
Wave Model of Light

- Physical Optics*
- Maxwell's Equations*
- Diffraction & Interference*



Particle Model of Light

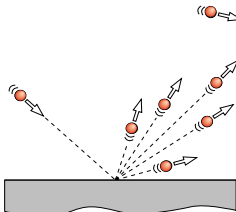
- Geometrical Optics*
- Phenomenological Description*
- Intensity and Color*



Radiative Transfer

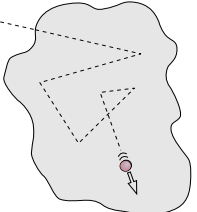
- Interaction of light with matter at the macroscopic scale
- Based on energy conservation
- Expressed by equation of transfer
- Originated in astrophysics
 - Schuster 1905*
 - Schwarzschild 1906*
 - Chandrasekhar 1947*

Surface Scattering



Interface between materials

Volume Scattering



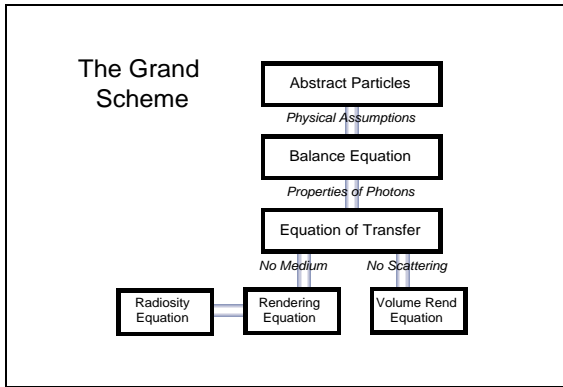
Participating medium

Work Involving Volume Interactions

- *Blinn 1982*
dusty surfaces
- *Kajiya and Von Herzen 1984*
clouds
- *Rushmeier and Torrance 1987, 1990*
radiosity + participating media
- *Krueger 1990*
volume rendering
- *He 1993*
subsurface scattering

Why study the equation of transfer?

- Describes important phenomena
 - Participating media*
 - Subsurface scattering*
- Provides a link to relevant literature
 - Radiative transfer*
 - Other fields*
- A useful conceptual model
 - Subsumes global illumination*
 - Clarifies assumptions & generalizations*

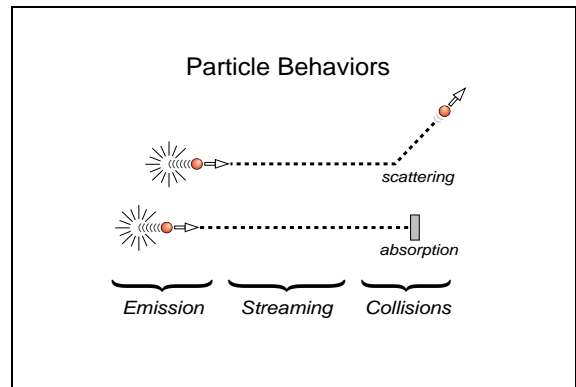


Related Phenomena

These very distinct physical phenomena have nearly identical governing equations:

- Radiative Transfer **Photons**
- Neutron Transport **Neutrons**
- Gas Dynamics **Molecules**

- ### Assumptions about Particles
- Collisions are instantaneous
 - Mono-energetic (gray, or one-speed)
 - No interactions among particles
 - Infinitely small & numerous
 - No external influences

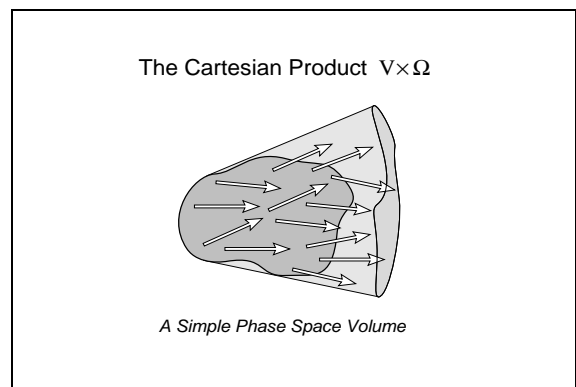
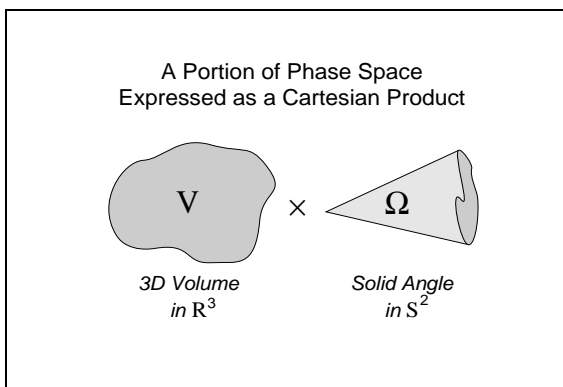


Deriving a Balance Equation

Particles are in constant motion, so equilibrium is dynamic. Changes due to all three behaviors must balance.

$$\left[\begin{array}{c} \text{change} \\ \text{due to} \\ \text{Emission} \end{array} \right] + \left[\begin{array}{c} \text{change} \\ \text{due to} \\ \text{Streaming} \end{array} \right] + \left[\begin{array}{c} \text{change} \\ \text{due to} \\ \text{Collisions} \end{array} \right] = 0$$

- ### Phase Space
- Used to represent state of particles
 - Captures all degrees of freedom
 - 3 for position
 - 2 for direction of travel
 - More for internal states (energy)
 - 5-dimensional in one-speed case
 - $R^3 \times S^2$ is the whole space



Phase Space Density $n(\mathbf{r}, \omega)$
(Angular Number Density)

$V \times \Omega$

Number of particles in box = $n(\mathbf{r}, \omega) d\mathbf{r} d\omega$

Phase Space Flux $\phi = v n$

$n ds dA d\omega$
 $= n (v dt) dA d\omega$
 $= \phi dt dA d\omega$

Flux from Density

$$\phi(\mathbf{r}, \omega) = v n(\mathbf{r}, \omega)$$

$\frac{1}{\text{m}^2 \text{sr s}}$ $\frac{\text{m}}{\text{s}}$ $\frac{1}{\text{m}^3 \text{sr}}$

A Balance Equation for Particles

$$\omega \cdot \nabla \phi(\mathbf{r}, \omega) + \sigma(\mathbf{r}) \phi(\mathbf{r}, \omega) = q(\mathbf{r}, \omega) + \int_{S^2} k(\mathbf{r}; \omega, \omega') \phi(\mathbf{r}, \omega') d\omega'$$

Particle conservation is expressed by an integro-differential equation

$\underbrace{\hspace{10em}}_f \quad \nabla$

Particle Behaviors

Emission Streaming Collisions

Emission Accounts for Creation of New Particles

$q(\mathbf{r}, \omega) \left[\frac{1}{\text{m}^3 \text{sr s}} \right]$

Phase space source function q accounts for new particles generated by chemical, thermal, or nuclear processes.

The Emission Term

$$\int_{\Omega} \int_V q(\mathbf{r}, \omega) d\mathbf{r} d\omega \rightarrow \left[\frac{1}{\text{s}} \right]$$

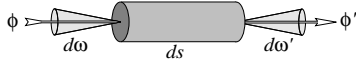
Particles per second

$\frac{1}{\text{m}^3 \text{sr s}}$ m^3 sr

Absorption Accounts for Particles that are Destroyed

The rate of absorption is proportional to the phase space flux ϕ .

Absorption:
Collisions per Unit Distance



The probability of absorption is proportional to the distance of travel.

The Absorption Term

$$\int_{\Omega} \int_V \sigma_a(\mathbf{r}) \phi(\mathbf{r}, \omega) d\mathbf{r} d\omega \rightarrow \left[\frac{1}{s} \right]$$

Particles per second

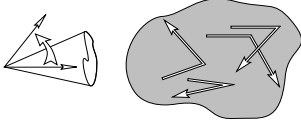
$\frac{1}{m}$

$\frac{1}{m^2 sr s}$

m^3

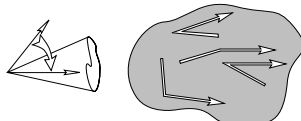
sr

Out-Scatter Accounts for
Particles Deflected Out of Ω



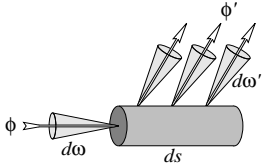
The rate of scattering is proportional to the phase space flux ϕ .

In-Scatter Accounts for
Particles Deflected Into Ω



Scattering collisions change the direction of travel instantaneously.

Scattering per Unit Distance
per Unit Solid Angle



The probability of scattering is proportional to the distance of travel and the solid angle.

The Scattering Terms

ω' for in-scatter
 ω for out-scatter

$$\int_{\Omega} \int_V \int_{S^2} k(\mathbf{r}; \omega \cdot \omega') \phi(\mathbf{r}, \omega) d\omega' d\mathbf{r} d\omega$$

$\frac{1}{m sr}$

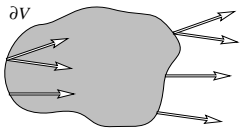
$\frac{1}{m^2 sr s}$

sr

m^3

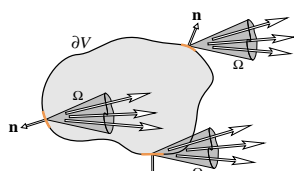
sr

Leakage Accounts for
Particles that Stream Out of V



Leakage depends only on the flux ϕ over ∂V

Computing Leakage Through ∂V



$$\int_{\Omega} \int_{\partial V} \phi(\mathbf{s}, \omega) \omega \cdot \mathbf{n}(\mathbf{s}) ds d\omega$$

Terms of the Balance Equation

Leakage

$$\omega \cdot \nabla \phi(\mathbf{r}, \omega)$$

Abs. & out-scatter

$$\sigma(\mathbf{r}) \phi(\mathbf{r}, \omega)$$

$$= \underbrace{q(\mathbf{r}, \omega)}_{\text{Emission}} + \underbrace{\int_{S^2} k(\mathbf{r}; \omega \cdot \omega') \phi(\mathbf{r}, \omega') d\omega'}_{\text{In-scatter}}$$

Boundary Conditions

- Needed for complete balance equation
The differential operator requires boundary conditions
- Two major forms
Explicit (Sources)
Implicit (Reflecting)
- Crucial in image synthesis
Surface effects predominate

The Semantics of Photons

- Particles carry energy
Phase space flux becomes radiance
 $L(\mathbf{r}, \omega) = h\nu \phi(\mathbf{r}, \omega)$
- Complex interactions with boundaries
Implicit boundary conditions essential
- Strong frequency dependence
Emission and collisions depend on ν

Radiance (Intensity)

$$L(\mathbf{r}, \omega) = h\nu \phi(\mathbf{r}, \omega) \rightarrow \left[\frac{\text{J}}{\text{m}^2 \text{sr s}} \right]$$

\uparrow
 $\left[\frac{\text{J}}{\text{m}^2 \text{sr s}} \right]$

\uparrow
 $\left[\frac{1}{\text{m}^2 \text{sr s}} \right]$

\downarrow
 $\left[\frac{\text{W}}{\text{m}^2 \text{sr}} \right]$

Directional Diffuse Reflection

$$dL(\mathbf{s}, \omega) = k_b(\mathbf{s}; \omega' \rightarrow \omega) L(\mathbf{s}, \omega') d\omega'$$

Boundary Conditions for the Equation of Transfer

$$L(\mathbf{s}, \omega) = \underbrace{\varepsilon_b(\mathbf{s}, \omega)}_{\text{surface emittance } \left[\frac{\text{W}}{\text{m}^2 \text{sr}} \right]} + \int_{S^2} \underbrace{k_b(\mathbf{s}; \omega' \rightarrow \omega)}_{\text{surface scattering kernel } \left[\frac{1}{\text{sr}} \right]} L(\mathbf{s}, \omega') d\omega'$$

Local Coordinate System of the BRDF

$$k_b(\mathbf{s}; \omega' \rightarrow \omega) = \rho_s(\theta_i, \phi_i, \theta_r, \phi_r) \cos \theta_i$$

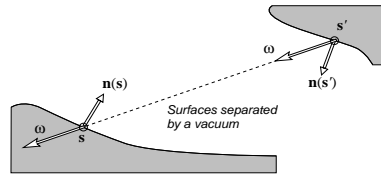
Vacuum Conditions

In a vacuum, the volume scattering and emission terms disappear, leaving only

$$\omega \cdot \nabla L(\mathbf{r}, \omega) = 0$$

which implies that radiance is constant along straight lines.

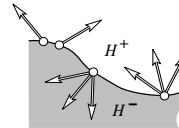
Radiance Determined from other Surfaces



The radiance at s in direction ω is given by the radiance at s' in direction ω .

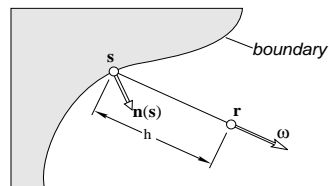
An Equation for Direct Exchange Between Surfaces

$$L(\mathbf{s}, \omega) = \epsilon_b(\mathbf{s}, \omega) + \int_{S^2} k_b(\mathbf{s}; \omega' \rightarrow \omega) L(\mathbf{s}', \omega') d\omega'$$



This is an equation for L defined on the set H^+ . The values of L on H^- are given by other surfaces.

Derivation of Integral Form



Integrate combined source term backwards along the ray until boundary is reached.

Non-Scattering Medium and Boundary

$$L(0) = v(x) \epsilon_b(x) + \int_0^x v(y) \epsilon(y) dy$$

$$e^{-\int_0^x \sigma(z) dz}$$

optical distance

