Chapter 2

Rendering Concepts

by Pat Hanrahan

2.1 Motivation

The progress in rendering in the last few years has been driven by a deeper and better understanding of the physics of materials and lighting. Physically based or realistic rendering can be viewed as the problem of simulating the propagation of light in an environment. In this view of rendering, there are sources that emit light energy into the environment; there are materials that scatter, reflect, refract, and absorb light; and there are cameras or retinas that record the quantity of light in different places. Given a specification of a scene consisting of the positions of objects, lights and the camera, as well as the shapes, material, and optical properties of objects, a rendering algorithm computes the distribution of light energy at various points in the simulated environment.

This model of rendering naturally leads to some questions, the answers to which form the subjects of this chapter.

1. What is light and how is it characterized and measured?

2. How is the spatial distribution of light energy described mathematically?

3. How does one characterize the reflection of light from a surface?

4. How does one formulate the conditions for the equilibrium flow of light in an environment?

In this chapter these questions are answered from both a physical and a mathematical point of view. Subsequent chapters will address specific representations, data structures, and algorithms for performing the required calculations by computer.

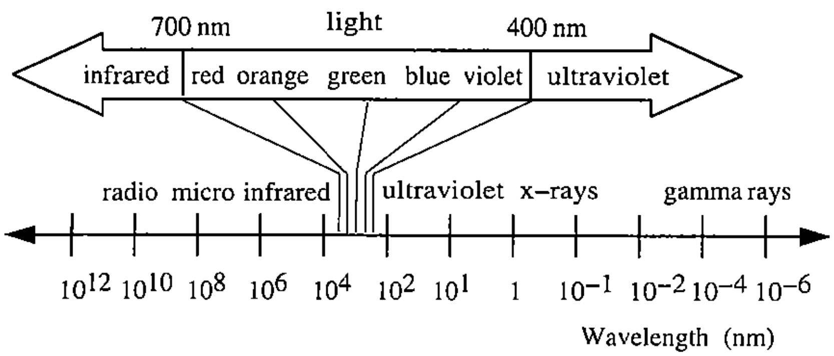


Figure 2.1: *Electromagnetic spectrum.*

2.2 Basic Optics

Light is a form of electromagnetic radiation, a sinusoidal wave formed by coupled electric and magnetic fields. The electric and magnetic fields are perpendicular to each other and to the direction of propagation. The frequency of the oscillation determines the wavelength. Electromagnetic radiation can exist at any wavelength. From long to short, there are radio waves, microwaves, infrared, light, ultraviolet, x-rays, and gamma rays (see Figure 2.1).

A pure source of light, such as that produced by a laser, consists of light at a single frequency. In the natural world, however, light almost always exists as a mixture of different wavelengths. Laser light is also *coherent*, that is, the source is tuned so that the wave stays in phase as it propagates. Natural light, in contrast, is *incoherent*.

Electromagnetic radiation can also be *polarized*. This refers to the preferential orientation of the electric and magnetic field vectors relative to the direction of propagation. Just as incoherent light consists of many waves that are summed with random phase, unpolarized light consists of many waves that are summed with random orientation. The polarization of the incident radiation is an important parameter affecting the reflection of light from a surface, but the discussion will be simplified by ignoring polarization.

The fact that light is just one form of electromagnetic radiation is of great benefit for computer graphics in that it points to theory and algorithms from many other disciplines, in particular, optics, but also more applied disciplines such as radar engineering and radiative heat transfer. The study of optics is typically divided into three subareas: geometrical or ray optics, physical or wave optics, and quantum or photon optics. Geometrical optics is most relevant to computer graphics since it focuses on calculating macroscopic properties of light

as it propagates through environments. Geometrical optics is useful to understand shadows, basic optical laws such as the laws of reflection and refraction, and the design of classical optical systems such as binoculars and eyeglasses. However, geometrical optics is not a complete theory of light. Physical or wave optics is necessary to understand the interaction of light with objects that have sizes comparable to the wavelength of the light. Physical optics allows us to understand the physics behind interference, dispersion, and technologies such as holograms. Finally, to explain in full detail the interaction of light with atoms and molecules quantum mechanics must be used. In the quantum mechanical model light is assumed to consist of particles, or photons. For the purposes of this book, geometrical optics will provide a full-enough view of the phenomena simulated with the radiosity methods.

2.3 Radiometry and Photometry

Radiometry is the science of the physical measurement of electromagnetic energy. Since all forms of energy in principle can be interconverted, a radiometric measurement is expressed in the SI units for energy or power, *joules* and *watts*, respectively. The amount of light at each wavelength can be measured with a spectroradiometer, and the resulting plot of the measurements is the spectrum of the source.

Photometry, on the other hand, is the psychophysical measurement of the visual sensation produced by the electromagnetic spectrum. Our eyes are only sensitive to the electromagnetic spectrum between the ultraviolet (380 nm) and the infrared (770 nm). The most prominent difference between two sources of light with different mixtures of wavelengths is that they appear to have different colors. However, an equally important feature is that different mixtures of light also can have different luminosities, or brightnesses.

Pierre Bouguer established the field of photometry in 1760 by asking a human observer to compare different light sources [35]. By comparing an unknown source with a standard source of known brightness—a candle at the time—the relative brightness of the two sources could be assessed. Bouguer's experiment was quite ingenious. He realized that *a human observer could not provide an accurate quantitative description of how much brighter one source was than another, but could reliably tell whether two sources were equally bright.*[1] Bouguer was also aware of the inverse square law. Just as Kepler and Newton had used it to describe the gravitational force from a point mass source, Bouguer reasoned that it also applied to a point light source. The experiment consisted of the

[1] This fact will be used in Chapter 9 when algorithms to select pixel values for display are examined.

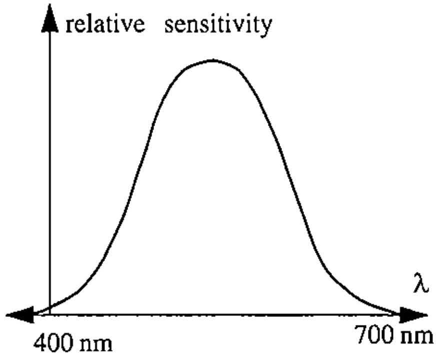


Figure 2.2: *Spectral luminous relative efficiency curve.*

observer moving the standard source until the brightnesses of the two sources were equal. By recording the relative distances of the two light sources from the eye, the relative brightnesses can be determined with the inverse square law.

Bouguer founded the field of photometry well before the mechanisms of human vision were understood. It is now known that different spectra have different brightnesses because the pigments in our photoreceptors have different sensitivities or responses toward different wavelengths. A plot of the relative sensitivity of the eye across the visible spectrum is shown in Figure 2.2; this curve is called the *spectral luminous relative efficiency curve*. The observer's response, R, to a spectrum is then the sum, or integral, of the response to each spectral band. This in turn is equal to the amount of energy at that wavelength, λ, times its relative luminosity.

$$R = \int_{380\text{nm}}^{770} V(\lambda)\, S(\lambda)\, d\lambda \tag{2.1}$$

where V is the relative efficiency and S is the spectral energy. Because there is wide variation between people's responses to different light sources, V has been standardized.

Radiometry is more fundamental than photometry, in that photometric quantities may be computed from spectroradiometric measurements. For this reason, it is best to use radiometric quantities for computer graphics and image synthesis. However, photometry preceded radiometry by over a hundred years, so much of radiometry is merely a modern interpretation of ideas from photometry.

As mentioned, the radiometric units for power and energy are the watt and joule, respectively. The photometric unit for luminous power is the *lumen*, and the photometric unit for luminous energy is the *talbot*. Our eye is most

sensitive to yellow-green light with a wavelength of approximately 555 nm that has a luminosity of 684 lumens per watt. Light of any other wavelength, and therefore any mixture of light, will yield fewer lumens per watt. The number of lumens per watt is a rough measure of the effective brightness of a light source. For example, the garden-variety 40-Watt incandescent light bulb is rated at only 490 lumens — roughly 12 lumens per watt. Of course, the wattage in this case is not the energy of the light produced, but rather the electrical energy consumed by the light bulb. It is not possible to convert electrical energy to radiant energy with 100% efficiency so some energy is lost to heat.

When we talk about light, power and energy usually may be used interchangeably, because the speed of light is so fast that it immediately attains equilibrium. Imagine turning on a light switch. The environment immediately switches from a steady state involving no light to a state in which it is bathed in light. There are situations, however, where energy must be used instead of power. For example, the response of a piece of film is proportional to the total energy received. The integral over time of power is called the *exposure*. The concept of exposure is familiar to anyone who has stayed in the sun too long and gotten a sunburn.

An important principle that must be obeyed by any physical system is the conservation of energy. This applies at two levels—a macro or global level, and a micro or local level.

- At the global level, the total power put into the system by the light sources must equal the power being absorbed by the surfaces. In this situation energy is being conserved. However, electrical energy is continuing to flow into the system to power the lights, and heat energy is flowing out of the system because the surfaces are heated.

- At the local level, the energy flowing into a region of space or onto a surface element must equal the energy flowing out. Accounting for all changes in the flow of light locally requires that energy is conserved. Thus, the amount of absorbed, reflected, and transmitted light must never be greater than the amount of incident light. The distribution of light can also become more concentrated or focused as it propagates. This leads to the next topic which is how to characterize the flow of light.

2.4 The Light Field

2.4.1 Transport Theory

The propagation of light in an environment is built around a core of basic ideas concerning the geometry of flows. In physics the study of how "stuff" flows

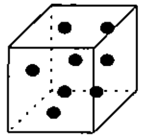


Figure 2.3: *Particles in a differential volume.*

is termed *transport theory*. The "stuff" can be mass, charge, energy, or light. Flow quantities are differential quantities that can be difficult to appreciate and manipulate comfortably. In this section all the important physical quantities associated with the flow of light in the environment will be introduced along with their application to computer graphics.

The easiest way to learn transport quantities is to think in terms of particles (think of photons). Particles are easy to visualize, easy to count, and therefore easy to track as they flow around the environment. The particle density $p(\mathbf{x})$ is the number of particles per unit volume at the point $\mathbf{x}$ (see Figure 2.3). Then the total number of particles, $P(\mathbf{x})$, in a small differential volume dV is

$$P(\mathbf{x}) = p(\mathbf{x})\, dV \tag{2.2}$$

Note that the particle density is an intrinsic or differential quantity, whereas the total number of particles is an absolute or extrinsic quantity.

Now imagine a stream of particles all moving with the same velocity vector $\vec{v}$; that is, if they are photons, not only are they all moving at the speed of light, but they are all moving in the same direction. We wish to count the total number of particles flowing across a small differential surface element dA in a slice of time dt. The surface element is purely fictitious and introduced for convenience and may or may not correspond to a real physical surface. In time dt each particle moves a distance $\vec{v}dt$. How many particles cross dA? This can be computed using the following observation: consider the tube formed by sweeping dA a distance $v\,dt$ in the direction $-\vec{v}$. All particles that cross dA between t and $t + dt$ must have initially been inside this tube at time t. If they were outside this tube, they would not be moving fast enough to make it to the surface element dA in the allotted time. This implies that one can compute the number of particles crossing the surface element by multiplying the particle volume density times the volume of the tube. The volume of the tube is just equal to its base (dA) times its height, which is equal to $v \cos\theta\, dt$. Therefore, as depicted in Figure 2.4, the total number of particles crossing the surface is

$$\begin{aligned} P(\mathbf{x}) &= p(\mathbf{x})\, dV \\ &= p(\mathbf{x})(v\, dt \cos\theta)\, dA \end{aligned} \tag{2.3}$$

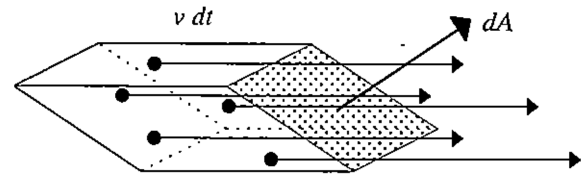


Figure 2.4: *Total particles crossing a surface.*

Note that the number of particles flowing through a surface element depends on both the area of the surface element and its orientation relative to the flow. Observe that the maximum flow through a surface of a fixed size occurs when the surface is oriented perpendicular to the direction of flow. Conversely, no particles flow across a surface when it is oriented parallel to the flow. More specifically, the above formula says that the flow across a surface depends on the cosine of the angle of incidence between the surface normal and the direction of the flow. This fact follows strictly from the geometry of the situation and does not depend on what is flowing.

The number of particles flowing is proportional both to the differential area of the surface element and to the interval of time used to tally the particle count. If either the area or the time interval is zero, the number of particles flowing is also zero and not of much interest. However, we can divide through by the time interval dt and the surface area dA and take the limit as these quantities go to zero. This quantity is called the *flux*.

More generally all the particles through a point will not be flowing with the same speed and in the same direction. Fortunately, the above calculation is fairly easy to generalize to account for a distribution of particles with different velocities moving in different directions. The particle density is now a function of two independent variables, position $\mathbf{x}$ and direction $\vec{\omega}$. Then, just as before, the number of particles flowing across a differential surface element in the direction $\vec{\omega}$ equals

$$P(\mathbf{x}, \vec{\omega}) = p(\mathbf{x}, \vec{\omega}) \cos\theta\, d\omega\, dA \tag{2.4}$$

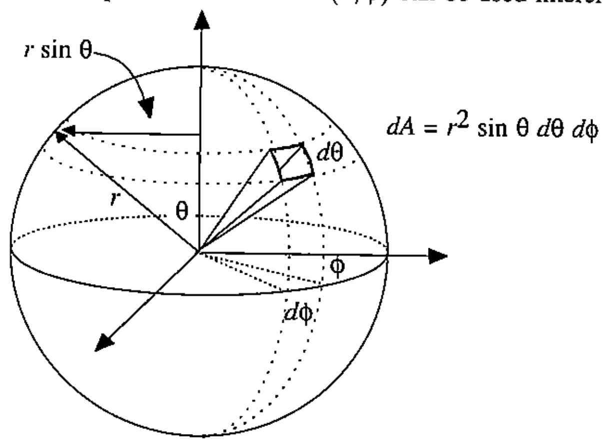
Here the notation $d\omega$ is introduced for the differential solid angle. The direction of this vector is in the direction of the flow, and its length is equal to the small differential solid angle of directions about $\vec{\omega}$. For those unfamiliar with solid angles and differential solid angles, please refer to the box.

2.4.2 Radiance and Luminance

The above theory can be immediately applied to light transport by considering light as photons. However, rendering systems almost never need consider (or at

Angles and Solid Angles

A direction is indicated by the vector $\vec{\omega}$. Since this is a unit vector, it can be represented by a point on the unit sphere. Positions on a sphere in turn can be represented by two angles: the number of degrees from the North Pole or zenith, θ, and the number of degrees about the equator or azimuth, ϕ. Directions $\vec{\omega}$ and spherical coordinates (θ, ϕ) can be used interchangeably.



A big advantage of thinking of directions as points on a sphere comes when considering differential distributions of directions. A differential distribution of directions can be represented by a small region on the unit sphere.

least have not considered up to this point) the quantum nature of light. Instead, when discussing light transport, the stuff that flows, or flux, is the radiant energy per unit time, or radiant power Φ, rather than the number of particles. The radiant energy per unit volume is simply the photon volume density times the energy of a single photon hc/λ, where h is Planck's constant and c is the speed of light. The radiometric term for this quantity is *radiance*.

$$L(\mathbf{x}, \vec{\omega}) = \int p(\mathbf{x}, \vec{\omega}, \lambda) \frac{hc}{\lambda} \, d\lambda \qquad (2.6)$$

Radiance is arguably the most important quantity in image synthesis. Defined precisely, radiance is power per unit projected area perpendicular to the ray per unit solid angle in the direction of the ray (see Figure 2.5). The definition in equation 2.6 is that proposed by Nicodemus [174], who was one of the first authors to recognize its fundamental nature.

The radiance distribution completely characterizes the distribution of light

The area of a small differential surface element on a sphere of radius r is

$$dA = (r\, d\theta)\,(r \sin\theta \, d\phi) = r^2 \sin\theta \, d\theta \, d\phi$$

Here $r\,d\theta$ is the length of the longitudinal arc generated as θ goes to $\theta + d\theta$. Similarly $r \sin\theta d\phi$ is the length of the latitudinal arc generated as ϕ goes to $\phi + d\phi$. The product of these two lengths is the differential area of that patch on the sphere.

This derivation uses the definition of angle in radians: the angle subtended by a circular arc of length l is equal to l/r. The circle itself subtends an angle of 2π radians because the circumference of the circle is $2\pi r$. By using a similar idea we can define a solid angle. The solid angle subtended by a spherical area a is equal to a/r^2. This quantity is the measure of the angle in *steradians* (radians squared), denoted sr. A sphere has a total area of $4\pi r^2$, so there are 4π steradians in a sphere.
A *differential solid angle*, indicated as $d\omega$, is then

$$d\omega = \frac{dA}{r^2} = \sin\theta \, d\theta \, d\phi \qquad (2.5)$$

It is very convenient to think of the differential solid angle as a vector, $d\vec{\omega}$. The direction of $d\vec{\omega}$ is in the direction of the point on the sphere, and the length of $d\vec{\omega}$ is equal to the size of the differential solid angle in that direction.

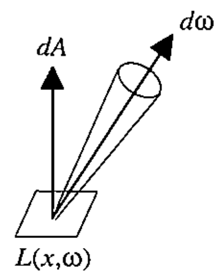


Figure 2.5: *The* radiance *is the power per unit projected area perpendicular to the ray per unit solid angle in the direction of the ray.*

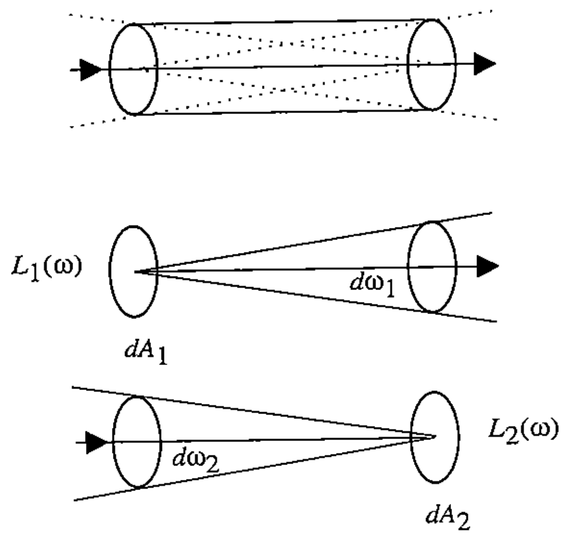


Figure 2.6: *Equality of flux leaving the first surface and flux arriving on the second surface.*

in a scene. Note that it is a function of five independent variables, three that specify position and two that specify direction. All other radiometric quantities can be computed from it. For example, the differential flux in a small beam with cross-sectional area dA and differential solid angle $d\omega$ is

$$d\Phi = L(\mathbf{x}, \vec{\omega}) \cos\theta \, d\omega \, dA \tag{2.7}$$

This follows immediately from the earlier discussion of particle transport.

To emphasize further the importance of radiance, consider the following two properties:

1. *The radiance in the direction of a light ray remains constant as it propagates along the ray* (assuming there are no losses due to absorption or scattering). This law follows from the conservation of energy within a thin pencil of light, as shown in Figure 2.6.

   The total flux leaving the first surface must equal the flux arriving on the second surface.

   $$L_1 \, d\omega_1 \, dA_1 = L_2 \, d\omega_2 \, dA_2 \tag{2.8}$$

   but $d\omega_1 = dA_2/r^2$ and $d\omega_2 = dA_1/r^2$, thus,

   $$T = d\omega_1 \, dA_1 = d\omega_2 \, dA_2 = \frac{dA_1 \, dA_2}{r^2} \tag{2.9}$$

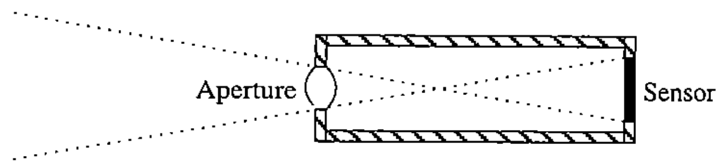


Figure 2.7: *A simple exposure meter.*

   This quantity T is called the *throughput* of the beam; the larger the throughput, the bigger the beam. This immediately leads to the conclusion that

   $$L_1 = L_2 \tag{2.10}$$

   and hence, the invariance of radiance along the direction of propagation. *As a consequence of this law, radiance is the numeric quantity that should be associated with a ray in a ray tracer.*

2. *The response of a sensor is proportional to the radiance of the surface visible to the sensor.*

   To prove this law, consider the simple exposure meter in Figure 2.7. This meter has a small sensor with area a and an aperture with area A. The response of the sensor is proportional to the total integrated flux falling on it.

   $$R = \int_A \int_\Omega L \cos\theta \, d\omega \, dA = LT \tag{2.11}$$

   Thus, assuming the radiance is constant in the field of view, the response is proportional to the radiance. The constant of proportionality is the throughput, which is only a function of the geometry of the sensor. The fact that the radiance at the sensor is the same as the radiance at the surface follows from the invariance of radiance along a ray.

   This law has a fairly intuitive explanation. Each sensor element sees that part of the environment inside the beam defined by the aperture and the receptive area of the sensor. If a surface is far away from the sensor, the sensor sees more of it. Paradoxically, one might conclude that the surface appears brighter because more energy arrives on the sensor. However, the sensor is also far from the surface, which means that the sensor subtends a smaller angle with respect to the surface. The increase in energy resulting from integrating over a larger surface area is exactly counterbalanced by the decrease in percentage of light that makes it to the sensor. This property of radiance explains why a large uniformly illuminated painted wall appears equally bright over a wide range of viewing distances.

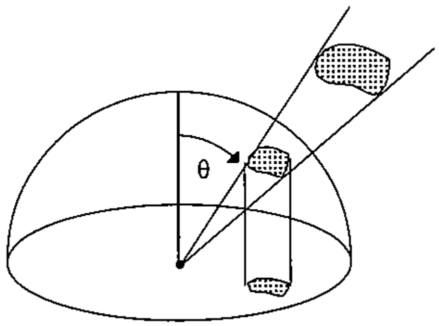


Figure 2.8: *Projection of differential area.*

As a consequence, the radiance from a surface to the eye is the quantity that should be output to the display device.

2.4.3 Irradiance and Illuminance

The two properties of radiance described in the previous section were derived by considering the total flux within a small beam of radiation. Another very important quantity is the total energy per unit area incident onto a surface with a fixed orientation. This can be computed by integrating the incident, or incoming radiance, L_i, over a hemisphere, Ω.

$$d\Phi = \left[\int_\Omega L_i \cos\theta \, d\omega\right] dA \tag{2.12}$$

The *irradiance*, E, is the radiant energy per unit area falling on a surface (the corresponding photometric quantity is the *illuminance*).

$$E \equiv \frac{d\Phi}{dA} \tag{2.13}$$

or

$$E = \int_\Omega L \cos\theta \, d\omega \tag{2.14}$$

The quantity $\cos\theta \, d\omega$ is often referred to as the *projected solid angle*. It can be thought of as the projection of a differential area on a sphere onto the base of the sphere, as shown in Figure 2.8.

This geometric construction shows that the integral of the projected solid angle over the hemisphere is just π, the area of the base of a hemisphere with unit radius. This result can also be derived directly by computing the following integral:

$$\begin{aligned}\int_\Omega \cos\theta \, d\omega &= \int_0^{2\pi}\int_0^{\pi} \cos\theta \, \sin\theta \, d\theta \, d\phi \\ &= -\int_0^{2\pi}\int_0^{\pi} \cos\theta \, d\cos\theta \, d\phi \\ &= -2\pi \frac{\cos^2\theta}{2}\bigg|_0^{\pi/2} \\ &= \pi \end{aligned} \tag{2.15}$$

Note that if all rays of light are parallel, which occurs if a single distant source irradiates a surface, then the integral reduces to the simple formula

$$E = E_0 \cos\theta \tag{2.16}$$

where E_0 is the energy per unit perpendicular area arriving from the distant source.

2.4.4 Radiosity and Luminosity

As the title of this book suggests, *radiosity* is another important quantity in image synthesis. Radiosity, B, is very similar to irradiance. Whereas irradiance is the energy per unit area incident onto a surface, radiosity is the energy per unit area that leaves a surface. It equals

$$B = \int_\Omega L_o \cos\theta \, d\omega \tag{2.17}$$

where L_o is the outgoing radiance.

The official term for radiosity is *radiant exitance*. Because of the widespread use of the term radiosity in the computer graphics literature, it will be used in this book. The photometric equivalent is *luminosity*.

2.4.5 Radiant and Luminous Intensity

Radiance is a very useful way of characterizing light transport between surface elements. Unfortunately, it is difficult to describe the energy distribution of a point light source with radiance because of the point singularity at the source. Fortunately, it is very easy to characterize the energy distribution by introducing another quantity—the *radiant* or *luminous intensity*.

Note that this use of "intensity" is very different from that typically used by the computer graphics community. Even more confusion results because intensity is often used to indicate radiance-like transport quantities in the physics community. The radiant intensity is quite similar to that used in the geometric optics community.

The energy distribution from a point light source expands outward from the center. A small beam is defined by a differential solid angle in a given direction. The flux in a small beam $d\omega$ is defined to be equal to

$$d\Phi \equiv I(\vec{\omega})\, d\omega \tag{2.18}$$

I is the radiant intensity of the point light source with units of power per unit solid angle. The equivalent photometric quantity is the luminous intensity.

The radiant intensity in a given direction is equal to the irradiance at a point on the unit sphere centered at the source. In the geometric optics literature intensity is defined to be the power per unit area (rather than per unit solid angle). In the case of a spherical wavefront emanating from a point source, the geometric optics definition is basically the same as the radiometric definition. However, in general, the wavefront emanating from a point source will be distorted after it reflects or refracts from other surfaces and so the definition in terms of solid angles is less general.

For an isotropic point light source,

$$I = \frac{\Phi}{4\pi} \tag{2.19}$$

Of course, a point source may act like a spotlight and radiate different amounts of light in different directions. The total energy emitted is then

$$\Phi = \int_{\Omega} I(\vec{\omega})\, d\omega \tag{2.20}$$

The irradiance on a differential surface due to a single point light source can be computed by calculating the solid angle subtended by the surface element from the point of view of the light source.

$$E = I\,\frac{d\omega}{dA} = \frac{\Phi}{4\pi}\,\frac{\cos\theta}{|\mathbf{x} - \mathbf{x}_s|^2} \tag{2.21}$$

where $|\mathbf{x} - \mathbf{x}_s|$ is the distance from the point to the surface element. Note the $1/r^2$ fall-off: this is the origin of the inverse square law.

The distribution of irradiance on a surface is often drawn using a contour plot or *iso-lux* diagram, while the directional distribution of the intensity from a point light source is expressed with a *goniometric* or *iso-candela* diagram.[2] This is a contour plot of equal candela levels as a function of the (θ, ϕ).

[2] See Chapter 10 for details of lighting specifications.

Physics	Radiometry	Radiometric Units
	Radiant energy	joules $[J = kg\, m^2/s^2]$
Flux	Radiant power	watts $[W = joules/s]$
Angular flux density	Radiance	$[W/m^2\, sr]$
Flux density	Irradiance	$[W/m^2]$
Flux density	Radiosity	$[W/m^2]$
	Radiant intensity	$[W/sr]$

Physics	Photometry	Photometric Units
	Luminous energy	talbot
Flux	Luminous power	lumens $[talbots/second]$
Angular flux density	Luminance	Nit $[lumens/m^2\, sr]$
Flux density	Illuminance	Lux $[lumens/m^2\, sr]$
Flux density	Luminosity	Lux $[lumens/m^2\, sr]$
	Luminous intensity	Candela $[lumens/sr]$

Table 2.1: *Radiometric and photometric quantities.*

2.4.6 Summary of Radiometric and Photometric Quantities

In most computer graphics systems, optical quantities are simply colors denoted by red, green, and blue triplets. These triplets are used to specify many quantities including light sources, material properties, and intermediate calculations.[3] As noted, there is a small but finite number (six to be exact) of radiometric (photometric) quantities that characterize the distribution of light in the environment. They are the radiant energy (luminous energy), radiant power (luminous power), radiance (luminance), irradiance (illuminance), radiosity (luminosity), and radiant intensity (luminous intensity). These quantities and their units are summarized in Table 2.1.

[3] A more complete treatment of color specification is given in Chapter 9.

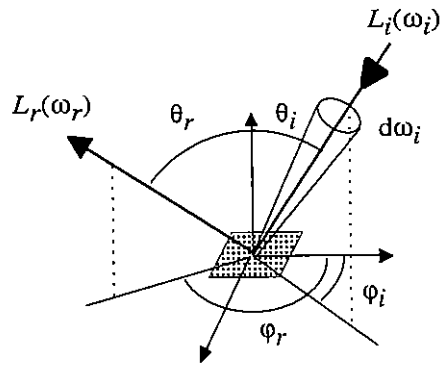


Figure 2.9: *Bidirectional reflection distribution function.*

2.5 Reflection Functions

The next question is how to characterize the reflection of light from a surface. Reflection is defined as the process by which light incident on a surface leaves that surface from the same side. Transmission, absorption, spectral and polarization effects, fluorescence, and phosphorescence are also important to consider in developing an accurate model of the interaction of light with materials, but will not be treated in detail here. Instead, this section will concentrate on nomenclature and the general properties that are satisfied by all reflection functions.

2.5.1 The Bidirectional Reflection Distribution Function

Consider the light incident on a surface from a small differential solid angle in the direction $\vec{\omega}_i$. The amount of reflected light in another direction $\vec{\omega}_r$ is proportional to the incident irradiance from $\vec{\omega}_i$ (see Figure 2.9). That is,

$$dL_r(\vec{\omega}_r) \propto dE(\vec{\omega}_i) \tag{2.22}$$

Equation 2.22 simply states that an increase in the incident light energy per unit area results in a corresponding increase in the reflected light energy. The incident irradiance can be increased by increasing either the solid angle subtended by the source or the energy density in the beam.

The constant of proportionality is termed the *bidirectional reflection distribution function*, or BRDF.

$$f_r(\vec{\omega}_i \to \vec{\omega}_r) \equiv \frac{L_r(\vec{\omega}_r)}{L_i(\vec{\omega}_i)\ \cos\theta_i\ d\omega_i} \tag{2.23}$$

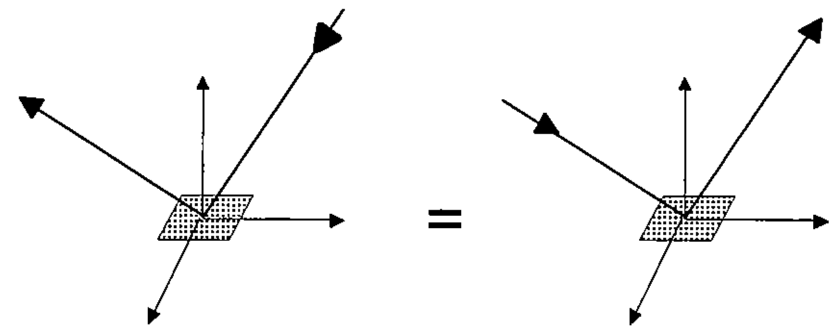


Figure 2.10: *Helmholtz reciprocity principle.*

More precisely, the BRDF is defined to be the ratio of the reflected radiance in the direction $\vec{\omega}_r$ to the differential irradiance from the incident direction $\vec{\omega}_i$ that produces it. The BRDF is bidirectional because it depends on two directions. Often, the dependence on the four angles is made explicit by writing the BRDF as $f_r(\theta_i, \phi_i; \theta_r, \phi_r)$. The BRDF is a distribution function because it is strictly positive. Since it gives the concentration of flux per steradian, it may take on any value between zero and infinity. The BRDF has units of inverse steradians.

The BRDF has several interesting properties:

1. If the BRDF is based on physical laws, then it will remain unchanged if the incident and reflected directions are interchanged. That is,

$$f_r(\vec{\omega}_r \to \vec{\omega}_i) = f_r(\vec{\omega}_i \to \vec{\omega}_r) \tag{2.24}$$

   This *Helmholtz reciprocity principle* is equivalent to saying that if a photon moves along a path, it will follow the same path if its direction is reversed (see Figure 2.10).

2. The BRDF is, in general, *anisotropic*. That is, if the incident and reflected directions are fixed and the underlying surface is rotated about the surface normal, the percentage of light reflected may change (see Figure 2.11). Examples of anisotropic materials are brushed aluminum or cloth [134].

   Many materials, however, are smooth and their reflectivity does not depend on the surface's orientation. Thus, their reflection functions do not change if the surface is rotated, and

$$f_r((\theta_i, \phi_i + \phi) \to (\theta_r, \phi_r + \phi)) = f_r((\theta_i, \phi_i) \to (\theta_r, \phi_r)) \tag{2.25}$$

   This implies that the reflection function has only three degrees of freedom instead of four.

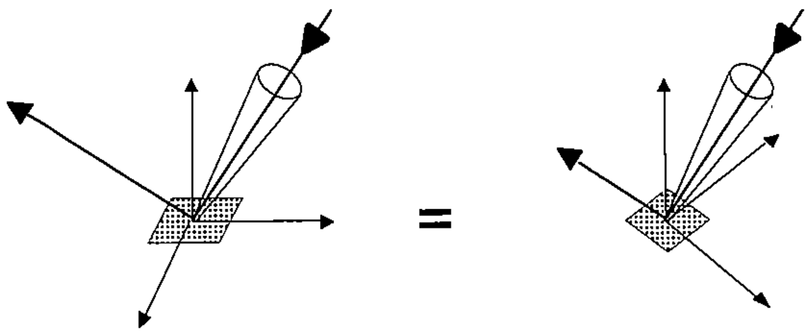


Figure 2.11: *The reflection may change with rotations of the surface due to anisotropy.*

Notice that adding light from another incident direction has no influence on the amount of light reflected from other incident directions. Thus, reflection behaves linearly, and hence the total amount of light reflected by a surface in a specific direction is given by a hemispherical integral over all possible incident directions. This leads to the *reflectance equation*:

$$L_r(\vec{\omega}_r) = \int_{\Omega_i} f_r(\vec{\omega}_i \to \vec{\omega}_r)\, L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i \tag{2.26}$$

Put another way, the reflected radiance in a particular direction is due to the radiance arriving from all directions weighted by the BRDF relating the incoming and reflected directions and by the projected solid angle.

2.5.2 Mirror Reflection

As an example of a BRDF, consider a perfect mirror and the geometry of the reflection. For a mirror, the angle of reflectance is equal to the angle of incidence, and the reflected vector is in the plane determined by the incident ray and surface normal vector. This implies that

$$\begin{aligned} \theta_r &= \theta_i \\ \phi_r &= \phi_i \pm \pi \end{aligned} \tag{2.27}$$

Second, consider the radiometry of reflection. For a mirror, the reflected radiance is exactly equal to the incident radiance.

$$L_r(\theta_r, \phi_r) = L_i(\theta_r, \phi_r \pm \pi) \tag{2.28}$$

This physical fact can be mathematically expressed with a BRDF involving delta functions.

$$f_{r,m} = \frac{\delta(\cos\theta_i - \cos\theta_r)}{\cos\theta_i}\,\delta(\phi_i - (\phi_r \pm \pi)) \tag{2.29}$$

Recall that the delta function has the following three properties:

1. $\delta(x) = 0 \quad \text{if } x \neq 0$
2. $\int_{-\infty}^{\infty} \delta(x)\, dx = 1$
3. $\int_{-\infty}^{\infty} \delta(x-y) f(x)\, dx = f(y)$

It can be verified that this leads to the correct reflected radiance by performing the hemispherical integral.

$$\begin{aligned} L_r(\theta_r, \phi_r) &= \int_{\Omega_i} \frac{\delta(\cos\theta_i - \cos\theta_r)}{\cos\theta_i}\,\delta(\phi_i - (\phi_r \pm \pi)) \\ &\quad \cdot L_i(\theta_i, \phi_i) \cos\theta_i\, d\theta_i\, d\phi_i \\ &= L_i(\theta_r, \phi_r \pm \pi) \end{aligned} \tag{2.30}$$

2.5.3 The Reflectance

Recall that the delta function can be interpreted as an infinitesimally thin, infinitely high spike with unit area. This implies that the BRDF, although always positive, may be infinite. Often it is more intuitive to work with a quantity that is bounded between 0 and 1. This quantity is called the *biconical reflectance*, or simply *reflectance*.

Consider the ratio of reflected flux to incident flux. Since the reflected flux must always be less than the incident flux giving rise to it, the reflectance must always be less than 1.

$$\begin{aligned} \frac{d\Phi_r}{d\Phi_i} &= \frac{\int_{\Omega_r} L_r(\vec{\omega}_r)\, \cos\theta_r\, d\omega_r}{\int_{\Omega_i} L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i} \\ &= \frac{\int_{\Omega_r}\int_{\Omega_i} f_r(\vec{\omega}_i \to \vec{\omega}_r)\, L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i\, \cos\theta_r\, d\omega_r}{\int_{\Omega_i} L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i} \end{aligned} \tag{2.31}$$

Unfortunately, the reflectance depends on the distribution of incoming light, L_i. If it is assumed that L_i is uniform and isotropic, then L_i can be taken out from the integral in both the numerator and the denominator. This results in the relationship between the reflectance and the BRDF which forms the definition of the reflectance:

$$\rho(\vec{\omega}_i \to \vec{\omega}_r) \equiv \frac{\int_{\Omega_r}\int_{\Omega_i} f_r(\vec{\omega}_i \to \vec{\omega}_r)\, \cos\theta_i\, d\omega_i\, \cos\theta_r\, d\omega_r}{\int_{\Omega_i} \cos\theta_i\, d\omega_i} \tag{2.32}$$

The reflectance involves a double integral over the incident and reflected directions for which the limits of integration have not yet been set. Three choices for the limits are a differential solid angle, a finite solid angle, or the entire hemisphere. Since this choice can be made for both the incident and the reflected directions, there are nine different reflectances. These are shown in table 2.2.

	$\vec{\omega}$	$\Delta\omega$	2π
$\vec{\omega}$	$\rho(\vec{\omega}_i \to \vec{\omega}_r)$	$\rho(\vec{\omega}_i \to \Delta\omega_r)$	$\rho(\vec{\omega}_i \to 2\pi)$
$\Delta\omega$	$\rho(\Delta\omega_i \to \vec{\omega}_r)$	$\rho(\Delta\omega_i \to \Delta\omega_r)$	$\rho(\Delta\omega_i \to 2\pi)$
2π	$\rho(2\pi \to \vec{\omega}_r)$	$\rho(2\pi \to \Delta\omega_r)$	$\rho(2\pi \to 2\pi)$

Table 2.2: *The nine biconical reflectances.*

The names of these reflectances are formed by combining the following words: *directional* (for differential solid angle), *conical* (for finite solid angle), and *hemispherical* (for a solid angle equal to the entire hemisphere). Thus, $\rho(\vec{\omega}_i \to \vec{\omega}_r)$, $\rho(\Delta\omega_i \to \Delta\omega_r)$, and $\rho(2\pi \to 2\pi)$ are referred to as the *bidirectional*, *biconical*, and *bihemispherical* reflectances, respectively. Perhaps the most interesting reflectance function is the *directional-hemispherical reflectance*, $\rho(\vec{\omega}_i \to 2\pi)$. This is the amount of light scattered into the entire hemisphere from a single incident direction. Since this quantity is the ratio of fluxes, it must be less than 1. However, be aware that this quantity can change with the angle of incidence.

2.5.4 Lambertian Diffuse Reflection

To illustrate the relationship between the BRDF and reflectance, consider the case of Lambertian diffuse reflectance. Diffuse reflectance is modeled by assuming that light is equally likely to be scattered in any direction, regardless of the incident direction. In other words, the BRDF is constant. Thus,

$$\begin{aligned} L_{r,d}(\vec{\omega}_r) &= \int_{\Omega_i} f_{r,d}\, L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i \\ &= f_{r,d} \int_{\Omega_i} L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i \\ &= f_{r,d} E \end{aligned} \tag{2.33}$$

This leads to two conclusions:

1. The value of the reflected radiance is proportional to the incident irradiance.

2. The reflected radiance is a constant and hence the same in all directions, since neither $f_{r,d}$ nor E depends on $\vec{\omega}_r$. This is true independent of the distribution of incoming light.

The fact that energy is conserved can be ensured by forcing the hemispherical–hemispherical reflectance to be less than 1.

$$\begin{aligned} \rho_d(2\pi \to 2\pi) = \frac{\Phi_{r,d}}{\Phi_i} &= \frac{\int_{\Omega_r} L_{r,d}(\vec{\omega}_r)\, \cos\theta_r\, d\omega_r}{\int_{\Omega_i} L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i} \\ &= \frac{L_{r,d}\int_{\Omega_r} \cos\theta_r\, d\omega_r}{E} \\ &= \frac{\pi L_{r,d}}{E} \\ &= \pi f_{r,d} \end{aligned} \tag{2.34}$$

It thus immediately follows that if the BRDF is a constant, then the reflectance is also a constant. More importantly, this relationship can be used to parameterize the BRDF in terms of the reflectance: $f_{r,d} = \rho_d/\pi$. Often, it is more intuitive to describe materials using their reflectances because they are constrained to lie between 0 and 1. Whenever a ρ is used in this text, it can safely be assumed to lie between 0 and 1.

Since the outgoing radiance is constant, the radiosity

$$B = \pi L_{r,d} \tag{2.35}$$

is related to the irradiance by the following equation:

$$\rho_d = \frac{B}{E} \tag{2.36}$$

Equation 2.36 states that for diffuse reflection, the reflectance is equal to the radiosity divided by the irradiance.

2.5.5 Glossy Reflection

In practice it is often convenient to treat the general BRDF as the sum of three qualitatively different components: mirror (or ideal) specular reflection, Lambertian (or ideal) diffuse reflection, and *glossy* reflection (see Figure 2.12). The diffuse and mirror reflection laws, Lambert's law, and the law of reflection, were discussed in the previous sections.

However, real materials are not perfectly diffuse or perfect mirror specular. This is to be expected since these models of reflection are the simplest mathematical abstractions of the properties of surfaces and materials. Real surfaces

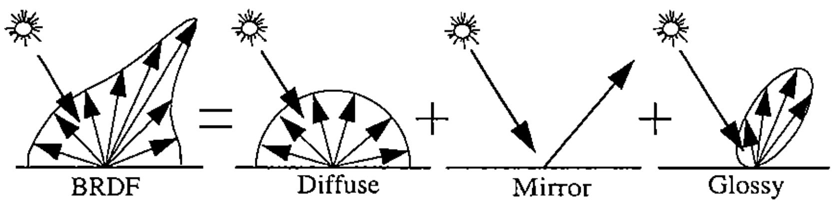


Figure 2.12: *Reflectance components.*

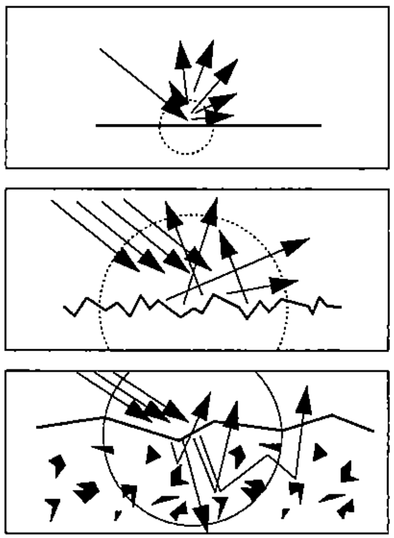


Figure 2.13: *Complex reflection distributions arise from rough surface and subsurface phenomena.*

are not planar and perfectly smooth and thus would not be expected to reflect light in just one direction. A real BRDF will thus contain a component between these limiting cases in which light hitting the surface from a certain direction is reflected into a complex distribution of outgoing directions.

The terminology for the various components is highly variable in the image synthesis literature. In particular, the intermediate component that we call glossy reflection is variously called specular, rough specular, wide and narrow diffuse, and directional diffuse. The term *glossy* has also been used in the surface reflection literature and has been selected instead for this work because its common usage is suggestive of the intended technical meaning.

Lord Rayleigh was the first to explain the effects of surface finish on the

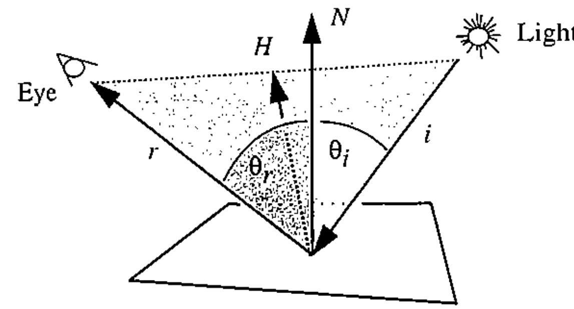


Figure 2.14: *Vectors for glossy reflection models.*

reflective properties of materials. He reasoned that a surface would become shinier if it were perfectly flat, or at least flat relative to the wavelength of the incident radiation. His theory is relatively easy to test because the wavelengths of common sources of radiation range from the macroscopic to the microscopic, and it can be verified that long wavelength radiation is more easily reflected from a surface. As shorter and shorter wavelengths are directed toward the surface, the ideal specular component decreases and the reflection becomes less focused. This transition occurs roughly when the wavelength of light becomes equal to the relative height changes in the surface. Thus, glossy reflection arises from the scattering of light from rough surfaces, an idea first proposed by Bouguer. The mirror specular term is considered to arise from perfectly smooth surfaces, while the Lambertian diffuse term arises from multiple surface reflections from very rough surfaces and from subsurface scattering (see Figure 2.13).

Another important optical effect is that glossy reflection increases at glancing angles of incidences and reflection. This is predicted by the Fresnel formula, which gives the relative percentage of light that is reflected or refracted across a planar boundary as a function of the angles of incidence and index of refraction.

In computer graphics glossy reflection from rough surfaces is typically modeled using the microfacet theory. This theory assumes the surface is made of little reflective facets, each behaving as a small mirror; that is, each reflecting light perfectly. This model predicts that the amount of light reflected from a light source toward the eye is equal to the relative number of microfacets oriented halfway between the eye and the light source. This model has been enhanced by many researchers [30, 65] and in its modern form consists of several terms

$$f_r = \frac{D\,G\,F}{4\cos\theta_r\,\cos\theta_i} \tag{2.37}$$

- D is the microfacet distribution. This distribution function gives the number of microfacets oriented in a particular direction. This function is typically modeled with the following formula (see Figure 2.14):

$$D(H, N, \kappa) = (N \cdot H)^{\kappa} \tag{2.38}$$

  Note that this distribution is maximal when H equals N, implying that the maximum number of microfacets are oriented parallel to the surface. Note also that κ controls the rate at which the distribution of microfacets falls off, and is related to the roughness of the surface.

- G is a geometric attenuation term accounting for self-shadowing. This arises because a rough surface is actually a height field, and facets in the valleys are less visible at glancing angles as facets at the peaks. This is an important effect, but very difficult to model precisely with a simple formula.

- F is the Fresnel reflection term related to a material's index of refraction.

The modeling of reflection of light from real materials is an interesting and important subject; however, space does not permit us to cover it in detail in this book. Models that can be found in the literature range from Phong's simple empirical model [181], to models of the form given above [30, 65, 236] that differ primarily in the details of the D function, to more recent (and complex) models such as that proposed by He *et al.* [118]. A good summary and description of the earlier models is given by Hall[114]. Subsurface reflection (see Figure 2.13) that has typically been modeled as part of the Lambertian diffuse component has also been reexamined to provide a more physically based model for biological materials such as skin and leaves [115].

2.6 The Rendering Equation

The reflectance equation makes it possible to compute the reflected light distribution from the incident light distribution and the BRDF of the material. The important remaining task is to specify, or preferably to compute, the incident light distribution. This is typically referred to as the *illumination model.*

The first and easiest case to consider is one with no occlusion and direct illumination from simple light sources. In this case there is typically a small number of point or distant lights, and it can be assumed that all light arrives at the surface; that is, there is no shadowing. Since this model does not consider the environment as a whole and only depends on the individual properties of the light sources and the surface being shaded, it is often called a *local* illumination

model. Shadows can be added by testing whether a point on the surface is visible to the light source. This is what is done in a ray tracer, but it requires access to the entire environment and is therefore an example of a *global* illumination model.

The second and considerably more difficult case is indirect illumination. In this case light may come from any surface in the environment, and it is very important to consider shadowing.

In the following sections, the interreflection of light between surfaces will be taken into account, and the *rendering equation* is derived from the reflectance equation. The *radiosity equation*, a simplified form of the rendering equation, that results by assuming all surfaces are Lambertian reflectors is also derived.

2.6.1 Local or Direct Illumination

It is easy to incorporate direct lighting from point light sources into the previous reflection models. Recall the reflectance equation

$$L_r(\vec{\omega}_r) = \int_{\Omega_i} f_r(\vec{\omega}_i \to \vec{\omega}_r)\, L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i \tag{2.39}$$

Recall also that the irradiance from a single point light source was derived,

$$E = \frac{\Phi}{4\pi} \frac{\cos\theta}{|\mathbf{x} - \mathbf{x}_s|^2} \tag{2.40}$$

If the direction to the light source is given by $\vec{\omega}_s$, then the radiance from a point light source can be expressed with a delta function.

$$L_i(\vec{\omega}_i) = \frac{\Phi}{4\pi\,|\mathbf{x} - \mathbf{x}_s|^2}\, \delta(\cos\theta_i - \cos\theta_s)\, \delta(\phi_i - \phi_s) \tag{2.41}$$

Substituting equation 2.41 into the reflectance equation yields

$$\begin{aligned} L_r(\vec{\omega}_r) &= \int f_r(\vec{\omega}_i \to \vec{\omega}_r)\, L_i(\vec{\omega}_i)\, \cos\theta_i\, d\omega_i \\ &= \frac{\Phi}{4\pi|\mathbf{x} - \mathbf{x}_s|^2} f_r(\vec{\omega}_r, \vec{\omega}_s)\, \cos\theta_s \end{aligned} \tag{2.42}$$

If there are n light sources, then the hemispherical integral collapses to a sum over the n sources. This is the lighting model used by 3D graphics workstations.

It is easy to extend this model to light sources with arbitrary directional distributions, as well as distant light sources. The above formulae are changed to use the radiant intensity in the direction of the surface. In principle, linear and area light sources can also be used, although this involves integrating the reflectance function over the range of possible directions incident from the light source. Nishita and Nakamae [175] and Amanatides [7] discuss this possibility.

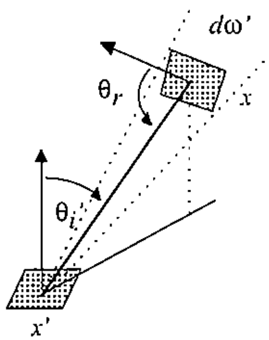


Figure 2.15: *Two point transport geometry.*

2.6.2 Global or Indirect Illumination

The first step in using a global illumination model is to relate the illumination on one surface to the reflected light distribution from another surface. This requires that the spatial dependence of radiance is made explicit and that occlusion is considered.

Using the fact that radiance is invariant along a ray, the incident radiance at $\mathbf{x}'$ due to the radiance from $\mathbf{x}$ is

$$L_i(\mathbf{x}', \vec{\omega}_i') = L_o(\mathbf{x}, \vec{\omega}_o) V(\mathbf{x}, \mathbf{x}') \tag{2.43}$$

where $\vec{\omega}_i$ is a direction vector from $\mathbf{x}'$ to $\mathbf{x}$, and $\vec{\omega}_o$ is in the opposite direction.

$$\vec{\omega}_i = -\vec{\omega}_o = \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|} \tag{2.44}$$

The function $V(\mathbf{x}, \mathbf{x}')$ is a visibility function. It is 1, if $\mathbf{x}$ and $\mathbf{x}'$ are mutually visible; otherwise it is 0.

Returning to the reflectance equation, the next step is to switch the hemispherical integral over all incident directions to an area integral over all the other surfaces in the environment. This is easily done by relating the solid angle subtended by the source to its projected surface area.

$$d\omega_i' = \frac{\cos\theta_o \, dA}{|\mathbf{x} - \mathbf{x}'|^2} \tag{2.45}$$

Dotting this to form the projected solid angle results in

$$d\omega_i' \cos\theta_o \, dA = G(\mathbf{x}, \mathbf{x}') \, dA \tag{2.46}$$

where

$$G(\mathbf{x}, \mathbf{x}') = G(\mathbf{x}', \mathbf{x}) = \frac{\cos\theta_i' \cos\theta_o}{|\mathbf{x} - \mathbf{x}'|^2} \tag{2.47}$$

Substituting $G(\mathbf{x}, \mathbf{x}')$ into the reflectance equation leads to the following integral equation over the surfaces, S:

$$L(\mathbf{x}', \vec{\omega}') = \int_S f_r(\mathbf{x}) \, L(\mathbf{x}, \vec{\omega}) \, G(\mathbf{x}, \mathbf{x}') \, V(\mathbf{x}, \mathbf{x}') \, dA \tag{2.48}$$

Since this equation only involves outgoing radiances and directions, the subscripts denoting incoming and outgoing directions can safely be dropped (except from f_r).

Equation 2.48 was first introduced to the computer graphics literature by Kajiya [135], who appropriately named it the *rendering equation*. Actually, his notation (and development) is slightly different than that used in equation 2.48. He introduced a new intensity quantity, $I(\mathbf{x} \to \mathbf{x}')$—the *two point transport intensity* from $\mathbf{x}$ to $\mathbf{x}'$ (see Figure 2.15). This intensity quantity is a function of surface position only and does not involve solid angles. The two point transport intensity is defined by the following equation:

$$d\Phi = I(\mathbf{x} \to \mathbf{x}') \, dA \, dA' = L(\mathbf{x}, \vec{\omega}) \, G(\mathbf{x}, \mathbf{x}') \, dA \, dA' \tag{2.49}$$

This is the flux flowing in the beam connecting dA to dA'. Equation 2.48 can be put in this form by multiplying both sides by $G(\mathbf{x}', \mathbf{x}'') \, dA' \, dA''$ which leads to the following equation:

$$I(\mathbf{x}' \to \mathbf{x}'') = G(\mathbf{x}', \mathbf{x}'') \int_S f_r(\mathbf{x} \to \mathbf{x}' \to \mathbf{x}'') \, V(\mathbf{x}, \mathbf{x}') I(\mathbf{x} \to \mathbf{x}') \, dA \tag{2.50}$$

Equation 2.50 defines the amount of light flowing from $\mathbf{x}$ to $\mathbf{x}'$ and reflected to $\mathbf{x}''$. Thus, it is sometimes referred to as the multipoint transport equation (see Figure 2.16). The quantity

$$f_r(\mathbf{x} \to \mathbf{x}' \to \mathbf{x}'') = f_r(\mathbf{x}', \vec{\omega}_i' \to \vec{\omega}_r') \tag{2.51}$$

is just a reparameterization of the BRDF.

There is one final step required to arrive at the full rendering equation, and that is to account for all modes of light transport at a surface. In an environment consisting only of opaque surfaces, the only other source of light is due to emission from the surface.

$$L(\mathbf{x}', \vec{\omega}') = L_e(\mathbf{x}', \vec{\omega}') + \int_S f_r(\mathbf{x}) \, L(\mathbf{x}, \vec{\omega}) \, G(\mathbf{x}, \mathbf{x}') \, V(\mathbf{x}, \mathbf{x}') \, dA \tag{2.52}$$

where L_e is the two point intensity of emitted light.

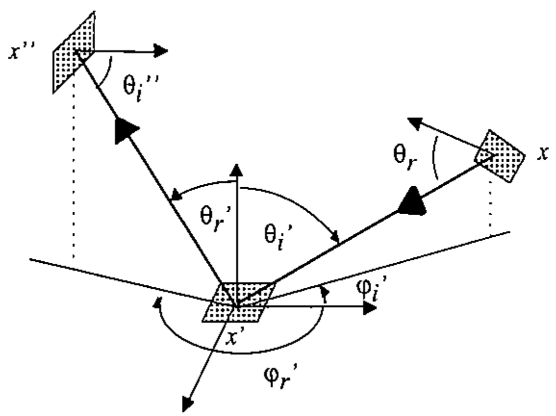


Figure 2.16: *Three point transport geometry.*

2.6.3 The Radiosity Equation

Finally, the rendering equation can be simplified given the *radiosity assumption*. In radiosity, it is assumed that all surfaces in the environment are Lambertian diffuse reflectors. Thus, the BRDF is independent of the incoming and outgoing directions and can be taken out from under the integral.

$$\begin{aligned} L(\mathbf{x}' \to \mathbf{x}'') &= L_e(\mathbf{x}' \to \mathbf{x}'') + f_r(\mathbf{x}') \int_S L(\mathbf{x} \to \mathbf{x}')\, G(\mathbf{x},\mathbf{x}')\, V(\mathbf{x},\mathbf{x}')\, dA \\ &= L_e(\mathbf{x}' \to \mathbf{x}'') + \frac{\rho(\mathbf{x}')}{\pi} \int_S L(\mathbf{x} \to \mathbf{x}')\, G(\mathbf{x},\mathbf{x}')\, V(\mathbf{x},\mathbf{x}')\, dA \end{aligned} \quad (2.53)$$

More importantly, the outgoing radiance from a Lambertian surface is the same in all directions and in fact equals the radiosity B divided by π. This leads to even more dramatic simplifications[4]

$$B(\mathbf{x}) = E(\mathbf{x}) + \rho(\mathbf{x}) \int_S B(\mathbf{x}')\, \frac{G(\mathbf{x},\mathbf{x}')\, V(\mathbf{x},\mathbf{x}')}{\pi}\, dA' \quad (2.54)$$

The rendering equation expresses the conservation of light energy at all points in space. The key feature of such an integral equation is that the quantity to be computed—in this case, the radiance or radiosity—appears on the left-hand side as well as under an integral on the right-hand side. For this reason, integral equations are notoriously difficult to solve. They very rarely have closed-form analytic solutions, and numerical methods must be used.

[4]Note the switch in notation: E is the energy per unit area emitted by the surface, or $\frac{L_e}{\pi}$. In addition, for clarity in the following chapters, the geometric term G(x,x') will absorb the visibility term and the π in the denominator.

Chapter 3

Discretizing the Radiosity Equation

3.1 The Radiosity Equation

The radiosity equation was derived at the end of Chapter 2 from the rendering equation under the assumption that all surfaces (and light sources) exhibit Lambertian diffuse reflection (emission). Repeating the radiosity equation 2.54:

$$B(\mathbf{x}) = E(\mathbf{x}) + \rho(\mathbf{x}) \int_S B(\mathbf{x}')\, G(\mathbf{x},\mathbf{x}')\, dA' \quad (3.1)$$

where the geometric term, $G(\mathbf{x},\mathbf{x}')$, now includes the visibility term, $V(\mathbf{x},\mathbf{x}')$, and division by π. (A complete table of the mathematical terms used in this chapter is provided in Tables 3.1 and 3.2.)

The radiosity, $B(\mathbf{x})$, describes an arbitrary *scalar* function across the surfaces (i.e., the radiosity function defines a single value at each location on a surface).[1] The potential complexity of the radiosity function is suggested by Figure 3.1, where the radiosity function across a partially shadowed polygon is plotted as a surface. The radiosity function is piecewise smooth, that is, it is continuous in all derivatives within regions bounded by discontinuities in value or derivatives. These characteristics will be discussed in much greater detail in chapters 6 and 8.

The dimension of the *function space* of the radiosity function, $B(\mathbf{x})$, is infinite (for a discussion of function spaces, refer to the box on page 45). This means that solving the radiosity equation for a point $\mathbf{x}$ on a surface does not

[1]A full solution to the radiosity problem must also take into account the distribution of energy across the visible spectrum (i.e., the *color* of the light). Assuming that the wavelength of light is not changed by interaction with surfaces (i.e., ignoring *fluorescence*), independent radiosity equations differing only in the reflectivities, ρ, can be formed and solved for each of a small number of wavelengths or color bands. The selection of these sample wavelengths and the reconstruction of colors suitable for display are discussed in Chapter 9. Elsewhere in the book, the radiosity problem will be discussed in terms of an achromatic (i.e., black, gray, white) world.