CS667 Lecture Notes: Scattering

Steve Marschner
Cornell University

3 September 2015

These notes are about how we describe optical scattering at surfaces and in volumes, radiometrically.

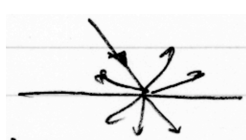
We've discussed how light travels through empty space. Today we look at how it interacts with stuff in the scene—after all, if the light just travels unimpeded the pictures will be pretty uninteresting.

1 Scattering

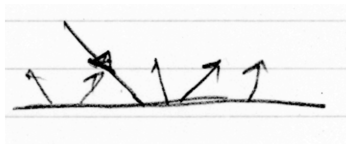
Scattering is when light coming from one direction ends up going into a range of directions. Specular reflection from smooth metal and glass surfaces is not scattering; it's reflection or refraction. I use the term "ideal specular" when I want to be clear I'm talking about this kind of one-direction-in, one-direction-out phenomenon.

There are three kinds of descriptions of scattering that are widely used in graphics:

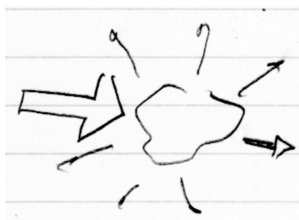
- BSDF, Bidirectional Scattering Distribution Function: describes scattering at infinitesimally thin surfaces (either the surface of a thick object or a thin object like a window or a sheet of paper). BSDF describes scattering both back to the same side (reflection, BRDF) and through to the other side (transmission, BTDF). This is a function of a point (on a surface) and two directions (incoming and outgoing, or incident and exitant).



- BSSRDF, Bidirectional Scattering-surface Reflectance Distribution Function: describes reflection from a thick object including light that scatters inside the material, not just what scatters at the surface. This means light can go in at one place and come out somewhere else, unlike with the BSDF. This is a function of two points (incoming and outgoing) as well as two directions.



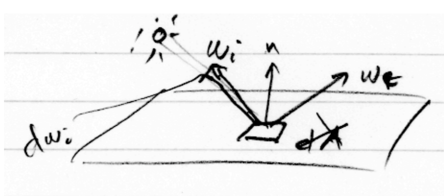
- Phase function: describes scattering that takes place over a 3D volume, not concentrated at a 2D surface. This is a function of a point (in space) and two directions.



In these notes I'll mostly be concerned with giving precise definitions of these quantities. In lecture I will discuss BSSRDF and the phase function when we encounter them later in the course.

2 BRDF

Think of a surface, with light incident on it at a particular point. Let's consider just the light arriving from an infinitesimal solid angle $d\omega_i$ around the incoming direction ω_i. If the radiance in this solid angle is L_i, then the (infinitesimal) irradiance on the surface is $dE_i(\omega_i) = L_i \mu(d\omega_i)$. (Remember that the projected solid angle measure μ has a factor of $\mathbf{n} \cdot \omega_i$ built into it, so that it assigns a smaller measure to more grazing solid angles.)



If this is a scattering surface (one that sends light to a range of directions, not just a single direction as a mirror would), this produces a distribution of infinitesimal reflected radiance $dL_r(\omega_r)$ over the hemisphere of outgoing directions ω_r.

The BRDF is the ratio of the reflected radiance to the incident irradiance:

$$f_s(\omega_i, \omega_r) = \frac{dL_r(\omega_r)}{dE_i(\omega_i)}$$

Now, if we want to get the (finite) radiance resulting from illumination from some distribution $L_i(\omega_i)$ we can just add up the contributions of illumination from differential solid angles all over the hemisphere—that is, we integrate:

$$L_r(\omega_r) = \int_{H^2} f_r(\omega_i, \omega_r) L_i(\omega_i) d\mu(\omega_i)$$

Operational version. If you like a concrete operational definition of this as a derivative, think of the experiment of illuminating the surface with a small area light source of radiance L_i (maybe a frosted light bulb with a dimmer to control L_i and an adjustable iris in front of it to control the solid angle Ω_i it illuminates) and measuring the reflected radiance L_r with a camera. The reflected radiance will be directly proportional to both the radiance of the source and, in the limit for small solid angles, the size of the solid angle. The BRDF is just the constant of proportionality between L_r and $L_i\mu(\Omega_i)$—it is the derivative of L_r with respect to E_i.

Mathematical version. If you like a more mathematical definition, you can think of light reflection as an operation on light distributions. You hand an incident radiance distribution $L_i : H^2 \to \mathbb{R}$ to the BRDF and it hands back a reflected distribution $L_r : H^2 \to \mathbb{R}$. Because of the superposition principle, this is a *linear* operator $\mathcal{R}$ on functions over the hemisphere:

Under reasonable conditions on the operator and the functions, this type of operator can always be expressed as an integral: the output is an integral of the input multiplied by a *kernel* function.

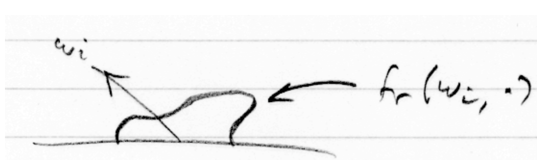
$$\begin{aligned} \mathcal{R} &: (H^2 \to \mathbb{R}) \to (H^2 \to \mathbb{R}) \\ &: L_i \mapsto \int_{H^2} f_r(\omega_i, \cdot) L_i(\omega_i) d\mu(\omega_i) \end{aligned}$$

We call this kernel function the BRDF.

2.1 BRDFs as densities

The BRDF is a function of four variables, which makes it a bit hard to think about sometimes. If we think of it in terms of one argument at a time, that can help.

For light arriving from ω_i, the BRDF $f_r(\omega_i, \cdot)$ is the density of reflectance over the outgoing hemisphere. Reflectance is a ratio of that tells the fraction of total irradiance reflected, and the BRDF describes the *distribution* of this reflectance over the hemisphere by giving the *density* function.



If we fix the outgoing direction instead, the function $f_r(\cdot, \omega_r)$ is another kind of density. It describes the distribution of something we might call "sensitivity" over the incoming hemisphere. By this I mean: the radiance in direction ω_r depends on light coming from many directions, and $f_r(\omega_i, \omega_r)$ tells you how sensitive it is to light from the particular direction ω_i.

2.2 Units of BRDF

One question that always comes up is, What does that mean, that the BRDF has units of inverse steradians? Why can't it just be unitless, since it relates radiance out to radiance in? I have three answers to this question.

First answer: I observed above that the BRDF is a density function that measures the density of reflectance (dimensionless ratio) over the hemisphere (measured in terms of solid angle). This is a density just like population density (people per square kilometer) or mass density in a solid (grams per cubic centimeter), so it has units of "reflectance per unit solid angle." But since reflectance is dimensionless we state this unit as just "per unit solid angle" or "one over steradians" (1/sr) or "inverse steradians" (sr^{-1}).

2.3 Properties of the BRDF

Not every function of two directions makes for a good BRDF. There are two properties all BRDFs have, physically: reflection conserves energy, and they obey Helmholtz reciprocity.

Energy conservation The basic requirement of energy conservation is that when a surface is illuminated with total irradiance E_i then the reflected radiant exitance M_r is less than E_i. This has to be true for all distributions of irradiance, so it has to be true of irradiance coming from a small solid angle Ω_i in the direction ω_i. If we integrate all the outgoing light for this case we have the radiant exitance:

$$\begin{aligned} L_r(\omega_r) &= f_r(\omega_r, \omega_i) L_i(\omega_i) \mu(\Omega_i) = f_r(\omega_r, \omega_i) E_i \\ M_r &= \int_{H^2} L_r(\omega_r)\, d\mu(\omega_r) \\ &= \int_{H^2} f_r(\omega_r, \omega_i)\, d\mu(\omega_r) \\ &= E_i \int_{H^2} f_r(\omega_r, \omega_i)\, d\mu(\omega_r). \end{aligned}$$

So radiant exitance is less than incident irradiance exactly when

$$\int_{H^2} f_r(\omega_r, \omega_i)\, d\mu(\omega_r) < 1.$$

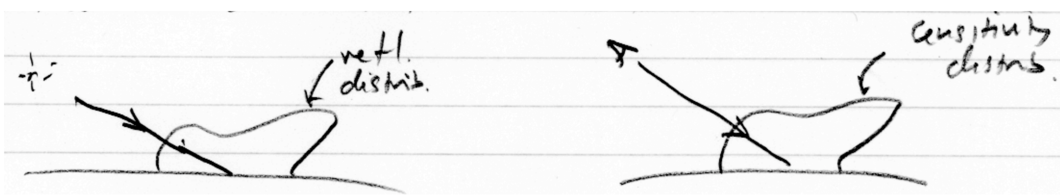
Here is a less hand-wavey proof that this guarantees energy conservation for arbitrary incident distributions:

$$
\begin{aligned}
L_r(\omega_r) &= \int_{H^2} f_r(\omega_r, \omega_i)\, L_i(\omega_i)\, d\mu(\omega_i) \\
M_r &= \int_{H^2} L_r(\omega_r)\, d\mu(\omega_r) \\
&= \int_{H^2} \int_{H^2} f_r(\omega_r, \omega_i)\, L_i(\omega_i)\, d\mu(\omega_i)\, d\mu(\omega_r) \\
&= \int_{H^2} L_i(\omega_i) \left[\int_{H^2} f_r(\omega_r, \omega_i)\, d\mu(\omega_r) \right] d\mu(\omega_i) \\
&< \int_{H^2} L_i(\omega_i)\, d\mu(\omega_i) = E_i
\end{aligned}
$$

Reciprocity This is part of a larger principle of reversibility of light transport paths, about which we'll hear more later in the course. In the context of the BRDF, the implication is that BRDFs are invariant with respect to swapping their arguments. That is:

$$f_r(\omega_1, \omega_2) = f_r(\omega_2, \omega_1).$$

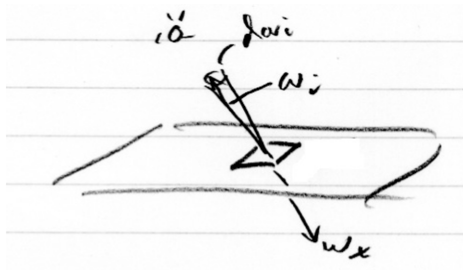
A physical interpretation is that the sensitivity distribution with the observer at a given position is the same as the reflected light distribution with the source at the same position.



This principle is true in reality (making it a great sanity-checking tool for measurements) and it is also a fundamental assumption of many advanced rendering algorithms. For this reason we try to ensure the BRDF models we invent are always reciprocal.

2.4 BTDF and BSDF

I have so far only talked in detail about the BRDF, but this is only half of the function—the BSDF—that I promised to talk about. The other half is the BTDF, and there is really nothing new at all: its definition is identical to the BRDF, but without the constraint that the two vectors are on the same side of the surface. The BTDF is radiance over irradiance, just like the BRDF.



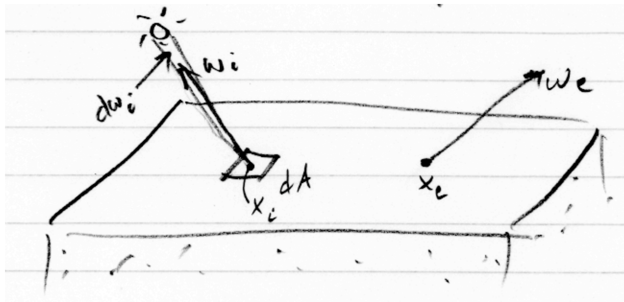
Nomenclature note: The R stands for “reflectance”; the T stands for “transmittance.” The “-ance” means “per unit input.” I’m not sure why we don’t tend to use the word “scatterance” for the BSDF, which is used in some other fields; we tend to let the S stand for “scattering.”

3 BSSRDF

The BSDF is for reflection from a surface—a perfectly thin surface—which has to happen at a single point. This is good for very thin sheets of material (a piece of paper) or for materials with a well-defined interface (metals, which are very opaque, or pure dielectrics, which are perfectly clear inside so that light only interacts at the surface). But not all materials are like this; many are thick and allow light to penetrate into the surface, then emerge some distance from the incident point. Some examples are marble (a crystalline material with lots of imperfections that scatter light), skin (a complex material with all kinds of microscopic structures), and milk (water with many protein molecules and fat droplets).

If we’d like to model these materials without actually including this subsurface light scattering in the rendering calculation, we need a reflectance distribution that accounts for scattering across the surface as well as scattering to different directions. This is the purpose of the *Bidirectional scattering-surface reflectance distribution function* or BSSRDF.

The BRDF can be considered a function of a point (the reflection point on the surface) and two directions (in and out directions). The difference between that and the BSSRDF is simply that the BSSRDF separates the single point into an incident point $\mathbf{x}_i$ and an exitant point $\mathbf{x}_r$. So the BSSRDF is a function, not just of a point and two directions (6 variables), but two points and two directions (8 variables).



As a result the incoming light needs to be integrated over the area of the surface:

$$L_r(\mathbf{x}_r, \omega_r) = \int_S \int_{H^2} S(\mathbf{x}_i, \omega_i, \mathbf{x}_r, \omega_r)\, L_i(\mathbf{x}_i, \omega_i)\, d\mu(\omega_i)\, dA(\mathbf{x}_i).$$

From this equation we can see that S has to have units of $\frac{1}{\mathrm{m}^2\mathrm{sr}}$, and we can interpret it as the derivative

$$S = \frac{dL_r(\mathbf{x}_r, \omega_r)}{d\Phi_i(\mathbf{x}_i, \omega_i)}$$

For a concrete interpretation of this derivative we can look at this experiment: Illuminate an area of surface S_i from solid angle Ω_i. This will result in some radiance being observed at a point $\mathbf{x}_r$ in the direction ω_r. Larger incident areas and solid angles will produce more radiance out for the same radiance in. The constant of proportionality between the radiance ratio L_r/L_i and the product $A(S_i)\mu(\Omega_i)$ is the BSSRDF.

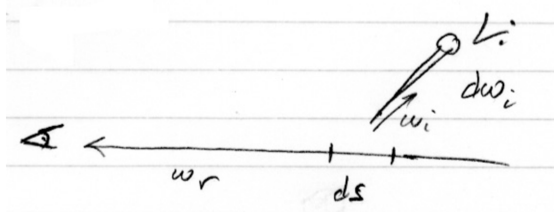
The BSSDRF has the same kind of reciprocity and energy conservation properties as the BRDF:

$$S(\mathbf{x}_i, \omega_i, \mathbf{x}_r, \omega_r) = S(\mathbf{x}_r, \omega_r, \mathbf{x}_i, \omega_i)$$
$$\int_S \int_{H^2} S(\mathbf{x}_i, \omega_i, \mathbf{x}_r, \omega_r)\, d\mu(\omega_r)\, dA(\mathbf{x}_r) < 1$$

4 Phase Function

Scattering can also happen in space, not at any particular surface. In this case scattering is distributed over a 3D volume, not just over a 2D surface.

Just like with the BRDF, we think of *single scattering* happening at a point. With the BRDF we had to integrate over solid angle to produce some reflected light; in the volume case we need to integrate over distance as well (to account for the difference between 2D and 3D). Scattering that contributes to radiance along a ray is distributed along the ray, rather than all happening at a single point where the ray intersects the reflecting surface. For this reason we need to look at a differential bit of ray in order to find the right description for volume scattering, and to get finite radiance we need to integrate along the ray and over the incident sphere (not hemisphere, since there is no surface) with respect to solid angle (not projected solid angle, since there is not surface, and therefore no surface normal).



$$dL_r(\mathbf{x}, \omega_r) = \tilde{f}_p(\omega_i, \omega_r)\, d\mu(\omega_i)\, ds$$
$$L_r(\mathbf{x}, \omega_r) = \int_0^\infty \int_{S^2} \tilde{f}_p(\mathbf{x}(s), \omega_i, \omega_r)\, L_i(\mathbf{x}(s), \omega_i\, d\sigma(\omega_i)\, ds$$

I'll call $\tilde{f}_p$ the "unnormalized phase function."

From the integral we can see $\tilde{f}_p$ has units of $\frac{1}{\mathrm{sr\,m}}$.

Unlike in surface reflection, it is customary to *normalize* phase functions. This means we scale the phase function so that its integral over the sphere is 1, and the resulting factored-out constant is called the scattering coefficient, σ_s. One reason this is done is that it's common for the (normalized) phase function to stay the same across space while the scattering coefficient changes.

$$\tilde{f}_p(\mathbf{x}, \omega_i, \omega_r) = \sigma_s(\mathbf{x})\, f_p(\omega_i, \omega_r)$$

I'll call f_p just the "phase function;" it has units of inverse steradians (1/sr). The scattering coefficient has units of inverse distance (1/m).

In many cases f_p is a property of the "stuff" that is floating around in the volume (smoke particles, water droplets, etc.) and σ_s is determined by how much of the stuff there is per unit volume.