Chapter 6

Reciprocity and Conservation Laws for General BSDF's

In this chapter, we derive a new reciprocity principle that holds for materials that transmit as well as reflect light.[1] According to this principle, the BSDF of any physically valid material must satisfy

$$\frac{f_{\mathrm{s}}(\omega_{\mathrm{i}} \to \omega_{\mathrm{o}})}{\eta_{\mathrm{o}}^2} \;=\; \frac{f_{\mathrm{s}}(\omega_{\mathrm{o}} \to \omega_{\mathrm{i}})}{\eta_{\mathrm{i}}^2}\,, \tag{6.1}$$

where η_{i} and η_{o} are the refractive indices of the materials containing ω_{i} and ω_{o} respectively. This is a generalization of the well-known condition for reflective materials, which states that the corresponding BRDF must be symmetric:

$$f_{\mathrm{r}}(\omega_{\mathrm{i}} \to \omega_{\mathrm{o}}) \;=\; f_{\mathrm{r}}(\omega_{\mathrm{o}} \to \omega_{\mathrm{i}})\,.$$

We also investigate how light scattering is constrained by the law of conservation of energy, and we derive a simple condition that must be satisfied by any BSDF that is energy-conserving.

These conditions are important for two reasons. First, they provide a convenient test of the plausibility of BSDF models in computer graphics. Second, they provide a minimal set

[1] A *reciprocity principle* is a statement that expresses some form of symmetry in the laws governing a physical system. Such principles have been proposed throughout physics and chemistry, and are often stated as a pair of hypothetical experiments whose outcomes are supposed to be the same.

of facts that can be assumed by any physically valid rendering system. Along these lines, in Chapter 7 we use the reciprocity condition mentioned above to derive a framework for which light, importance, and particles all obey the same transport rules (for any physically valid scene). This can make many rendering algorithms significantly easier to implement.

The main goal of this chapter is to derive the new reciprocity principle (6.1). Given an arbitrary material, we analyze its scattering properties when it is placed within an *isothermal enclosure* (i.e. one where all objects have the same temperature, and no heat is lost to the external environment). For such a system, which is said to be in *thermodynamic equilibrium*, the exchange of light energy between various parts of the enclosure is highly constrained by the laws of thermodynamics. This allows us to derive the reciprocity condition (6.1) from only two basic principles, namely *Kirchhoff's equilibrium radiance law*, and the *principle of detailed balance*. Note that even though our analysis takes place within an isothermal enclosure, the resulting reciprocity principle is valid generally (since the BSDF of a material is an inherent property).

We also discuss the historical origins of reciprocity principles. One of the first physicists to study these ideas was Helmholtz, who proposed a famous principle concerning the propagation of light through an optical system. However, it is important to note that Helmholtz himself did not make any statement that would imply the symmetry of BRDF's. As we will see, his reciprocity principle only applies to reflection from mirrors (rather than arbitrary materials), and thus it does not have any direct implications for the symmetry of general BRDF's.

We also discuss the subtleties that arise in rigorously justifying such principles. For example, we explain why the symmetry of BRDF's cannot be derived directly from the second law of thermodynamics, or from the principle of *time reversal invariance*.

This chapter is organized as follows. Section 6.1 describes the second law of thermodynamics, Kirchhoff's laws, and the principle of detailed balance. Section 6.2 shows how these ideas can be put together in a "thought experiment" to prove the desired reciprocity condition (6.1). Section 6.3 derives a separate condition to ensure that BSDF's are energy-conserving.

In the appendices we examine the history of reciprocity principles, and also their limitations. Appendix 6.A describes the Helmholtz reciprocity principle, and explains why it does not have any implications regarding the symmetry of BRDF's. Appendix 6.B describes a different reciprocity principle due to Lord Rayleigh, who appears to be the first person to state a principle for reflection from arbitrary surfaces. Appendix 6.C considers the principle of time reversal invariance, and explains why observable light scattering processes are *irreversible* in general. Finally, Appendix 6.D investigates the limitations of these reciprocity principles, by describing two situations in which they fail: namely, in the presence of absorbing media or external magnetic fields.

6.1 Thermodynamics, Kirchhoff's laws, and detailed balance

Consider an enclosure containing various kinds of matter, which is completely insulated from its surrounding environment. Eventually, the contents will reach a uniform temperature, and the system is said to be in *thermodynamic equilibrium*. At equilibrium, each portion of matter will be emitting, scattering, and absorbing energy at various wavelengths (e.g. in the thermal, visible, and ultraviolet portions of the spectrum), in a manner that depends on both the local material properties and the surrounding radiation field. Thus, energy is constantly being exchanged among different regions of the enclosure, but in such a way that the temperature everywhere remains constant.

In this section, we explain two basic facts about systems in thermodynamic equilibrium, which will be used to derive the reciprocity condition (6.1). These facts are:

1. The radiance in an isothermal enclosure is *uniform*, i.e. it is the same for all positions and directions. More precisely, it depends only on the temperature of the enclosure and the local refractive index, such that
$$\frac{L_\nu(\mathbf{x}, \omega, \nu)}{\eta(\mathbf{x}, \omega, \nu)^2}$$
is constant throughout the enclosure (this is called *Kirchhoff's equilibrium radiance law*). Here $(\mathbf{x}, \omega)$ is a ray, ν is a frequency, $L_\nu(\mathbf{x}, \omega, \nu)$ is the spectral radiance for this ray and frequency, and $\eta(\mathbf{x}, \omega, \nu)$ is the refractive index of the medium that surrounds

this ray. (The reason that η is a function of ω is to handle the case when $\mathbf{x}$ is on the boundary between two different media.)

2. For every process that transfers energy from one part of an isothermal system to another, there is a reverse process that transfers energy at the same rate in the opposite direction. This is known as the *principle of detailed balance*. For example, this principle states that for a system in thermodynamic equilibrium, the rates of emission and absorption for any given surface are equal.

We now explain these concepts in more detail. We first discuss the second law of thermodynamics, followed by Kirchhoff's equilibrium radiance law, and finally the principle of detailed balance. Our discussion of these ideas is based mainly on the excellent summary of [Milne 1930]; more detailed information can be found in Drude [1900], Siegel & Howell [1992], and [de Groot & Mazur 1962].

The second law of thermodynamics. Using the second law of thermodynamics, it is possible to derive important facts about the distribution of light energy in an isothermal enclosure. According to this principle, no ideal experiment can produce a temperature difference within the enclosure unless the experiment does work or modifies the external environment. For example, suppose that we divide the enclosure into two compartments separated by a surface S. Furthermore, suppose that S is transparent to light in a particular frequency band $[\nu_1, \nu_2]$, but reflects light at all other frequencies. Then by the second law, the rate of energy flow across this surface must be the same in both directions. Otherwise, the net flow would produce a temperature difference between the two sides of S, which could then be used to perform work.

The second law can be stated more precisely in terms of *entropy*. Entropy measures the amount of energy that can be transferred from one system to another, in the form of work. For a given system with a fixed energy, the entropy can range from zero to some maximum: if it is zero, then all of the energy in the system can be converted into work; while if it is at a maximum, then no work can be done at all. With respect to this concept, the second law states that the entropy of a closed, insulated system can never decrease, unless work is performed on it from some external source (see [de Groot & Mazur 1962, p. 20] for further details). Thus, once a system has reached thermodynamic equilibrium (the state of maximum entropy), it will remain in equilibrium, even if we perform ideal experiments such as adding barriers or mirrors, changing the locations of objects, etc.

Kirchhoff's laws. By proposing ideal experiments of this kind, Gustav Kirchoff was able to derive many interesting facts about the radiation in an isothermal enclosure [Milne 1930, p. 79], which are collectively known as *Kirchhoff's laws*.[2] Of these facts, we will need only his *equilibrium radiance law* mentioned above, which states that the quantity

$$\frac{L_\nu(\mathbf{x}, \omega, \nu)}{\eta(\mathbf{x}, \omega, \nu)^2} \tag{6.2}$$

is constant throughout the enclosure.[3] For example, if the objects in the enclosure are surrounded by a single medium, such as air, then this law states that the observed spectral radiance for all positions and directions will be the same. This is true even though the objects within the enclosure may have very different emission, scattering, and absorption properties. In fact, the observed spectral radiance depends only on temperature; given any two enclosures with different contents, but at the same temperature, the observed spectral radiance in these enclosures will be the same.

If the objects in the enclosure are surrounded by several different media, the spectral radiance will be proportional to η^2, as indicated by equation (6.2). This is one of the key facts that we will need to derive the reciprocity condition for general BSDF's.

Detailed balance. The other fact we need is the principle of detailed balance, which asserts that for a system in thermodynamic equilibrium, every detailed process that we choose to consider has a reverse process, and that the rates of these two processes are equal [van de Hulst 1980, p. 17]. For example, this principle asserts that in an isothermal enclosure, the

[2] In the heat transfer literature, *Kirchhoff's law* generally refers to one of these facts in particular, namely that the emissivity and absorptivity of real materials are the same [Siegel & Howell 1992, p. 66]. This was derived by Kirchhoff as a consequence of his equilibrium radiance law. Note that these results are not related to Kirchhoff's laws for electric circuits, which he proposed much earlier in 1845. Also note that the invariance of L/η^2 is often falsely attributed to Clausius (cf. Drude [1900, p.504]).

[3] Strictly speaking, this law is true only when some material in the enclosure is capable of emitting or absorbing radiation at the given frequency ν [Milne 1930, p. 80]. We can ensure that this is always true by assuming that the enclosure contains a *black body* (which absorbs and emits radiation at all frequencies).

emission and absorption rates of every surface are equal. This principle also applies to scattering, as we will see in the next section.

Detailed balance has been formulated and proven for any system that possesses time reversal invariance, both in classical systems and in quantum mechanics [de Groot & Mazur 1962], [van Kampen 1954], [Wigner 1954]. *Time reversal invariance* is one of the basic principles of physics, which states if the time variable is negated in all formulas and equations, then the laws of physics at their most microscopic level are unchanged (see Appendix 6.C). The only significant restriction of detailed balance is that for it to be valid, there must not be any external magnetic fields [de Groot 1963].

6.2 A reciprocity principle for general BSDF's

By combining Kirchhoff's equilibrium radiance law with the principle of detailed balance, we derive a reciprocity principle that holds for arbitrary physically valid materials. We also show that this principle cannot be derived from the second law alone.

Consider a small area $dA(\mathbf{x})$ within an isothermal enclosure (see Figure 6.1). We assume that $\mathbf{x}$ lies either on an opaque surface, or on the boundary between two non-absorbing media. We also assume that no external magnetic fields are present, so that the principle of detailed balance applies.

Consider the light that arrives from a small cone of directions $d\sigma(\omega_\mathrm{i})$, and is scattered toward another cone $d\sigma(\omega_\mathrm{o})$, where ω_i and ω_o can lie on either side of the surface. The scattering can be of any type: reflection or transmission, specular or non-specular. According to the definition of the BSDF (3.11), the power scattered from ω_i to ω_o is

$$\begin{aligned} d\Phi_1 &= L_\mathrm{o}(\omega_\mathrm{o})\, dA(\mathbf{x})\, d\sigma^\perp(\omega_\mathrm{o}) \\ &= L_\mathrm{i}(\omega_\mathrm{i})\, d\sigma^\perp(\omega_\mathrm{i})\, f_\mathrm{s}(\omega_\mathrm{i} \!\to\! \omega_\mathrm{o})\, dA(\mathbf{x})\, d\sigma^\perp(\omega_\mathrm{o})\,. \end{aligned}$$

On the other hand, the power scattered from ω_o to ω_i is

$$d\Phi_2 = L_\mathrm{i}(\omega_\mathrm{o})\, d\sigma^\perp(\omega_\mathrm{o})\, f_\mathrm{s}(\omega_\mathrm{o} \!\to\! \omega_\mathrm{i})\, dA(\mathbf{x})\, d\sigma^\perp(\omega_\mathrm{i})\,.$$

By the principle of detailed balance, the rates of scattering in these two directions are equal.

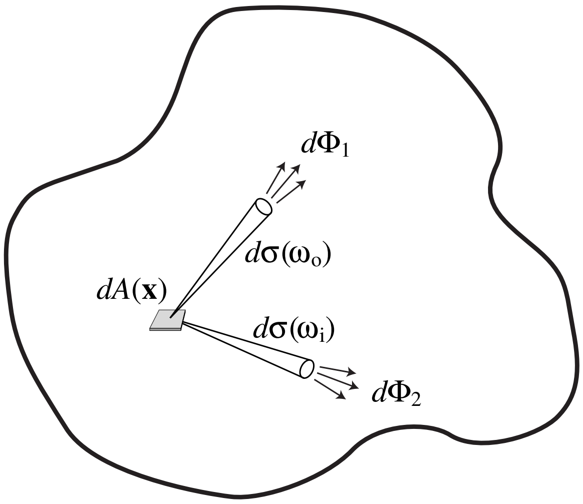


Figure 6.1: To prove a reciprocity condition for general BSDF's, we consider the light energy scattered between two directions ω_i and ω_o at a point $\mathbf{x}$ in an isothermal enclosure. By the principle of detailed balance, the rates of scattering from ω_i to ω_o and from ω_o to ω_i are equal ($d\Phi_1 = d\Phi_2$), while by Kirchhoff's equilibrium radiance law, the incident radiance from each direction is proportional to the refractive index squared. Putting these facts together, we get the desired reciprocity condition (6.1).

Thus we have $d\Phi_1 = d\Phi_2$, so that

$$L_\mathrm{i}(\omega_\mathrm{i})\, f_\mathrm{s}(\omega_\mathrm{i} \!\to\! \omega_\mathrm{o}) \;=\; L_\mathrm{i}(\omega_\mathrm{o})\, f_\mathrm{s}(\omega_\mathrm{o} \!\to\! \omega_\mathrm{i})\,.$$

Next, we consider the incident radiance values, $L_\mathrm{i}(\omega_\mathrm{i})$ and $L_\mathrm{i}(\omega_\mathrm{o})$. According to Kirchhoff's equilibrium radiance law, L_i/η^2 is constant throughout the enclosure, so that

$$\frac{L_\mathrm{i}(\omega_\mathrm{i})}{\eta_\mathrm{i}^2} \;=\; \frac{L_\mathrm{i}(\omega_\mathrm{o})}{\eta_\mathrm{o}^2}\,.$$

Putting these two facts together, we get the following result for physically valid BSDF's:

Theorem 6.1. *Let f_s be the BSDF for a physically valid surface, which is either the boundary of an opaque object or the interface between two non-absorbing media. Provided that there are no external magnetic fields, then*

$$\frac{f_\mathrm{s}(\omega_\mathrm{i} \!\to\! \omega_\mathrm{o})}{\eta_\mathrm{o}^2} \;=\; \frac{f_\mathrm{s}(\omega_\mathrm{o} \!\to\! \omega_\mathrm{i})}{\eta_\mathrm{i}^2}\,, \tag{6.3}$$

where $\eta_o = \eta(\omega_o)$ *is a function of* ω_o*, and similarly for* η_i. ■

This condition is clearly a generalization of the usual symmetry condition for BRDF's. The most significant change concerns BSDF's that describe the interface between two different refractive media. For this case, the ratio of f_s to f_s^* is $(\eta_o/\eta_i)^2$, so that radiance and importance are scaled differently when they are transmitted through the interface. This law is not limited to perfect specular refraction; it also includes diffusely transmitting materials, such as frosted glass.

Note that although this relationship was derived in an isothermal enclosure, it is valid in general. The BSDF is an inherent property of the surface, and does not change simply because the surrounding environment is isothermal. Also note that the enclosure does not have to be at a high temperature for this argument to hold, since even at ordinary temperatures, there is a small amount of thermal radiation in the visible wavelengths.

Insufficiency of the second law. Returning to the simpler case of opaque materials (BRDF's), it is sometimes claimed that the reciprocity condition for these materials can be derived directly from the second law of thermodynamics.[4] We show that this is false, by giving an example of a BRDF which is not symmetric, but where this lack of symmetry cannot be detected by any ideal experiment in an isothermal enclosure.

We consider a hypothetical surface that is similar to a mirror. For an ordinary mirror, light is reflected from the incident direction ω_i to the mirror direction ω_o, where the mirror direction is obtained by rotating ω_i by 180 degrees about the surface normal. We consider a new BRDF that modifies this rule: the mirror vector is obtained by rotating the incident vector by only 90 degrees about the surface normal, in a clockwise direction. Clearly, this new BRDF is not symmetric.

However, the new BRDF and the original mirror BRDF are *indistinguishable* in an isothermal enclosure. The reason is that the incident radiance is guaranteed to be uniform, and both of these BRDF's will map a uniform incident radiance function into a uniform exitant radiance function. Thus, there is no ideal experiment in an isothermal enclosure that

[4] For example, the BRDF reciprocity argument of Siegel & Howell [1992, p. 73] appears to depend only on the second law. However, their argument is flawed. To fix it, they require the principle of detailed balance; in which case their proof could be simplified by neglecting the transport path labeled $dA_1\, dA_3$ in their figure.

can distinguish between these two situations. This is why the principle of detailed balance is necessary; it provides more detailed information about how the incident and exitant radiance functions are related, by considering the energies traveling in opposite directions along the same path.

Reciprocity for spectral radiance. To be precise, the reciprocity condition (6.3) applies to spectral radiance (rather than radiance), and in fact it applies only to spectral radiance that is measured with respect to frequency (L_ν). When spectral radiance is measured with respect to wavelength (L_λ), this condition must be modified, since light undergoes a change in wavelength when it is transmitted into a different medium. In particular, the incident and transmitted wavelengths are related by

$$\lambda_{\mathrm{t}} \;=\; (\eta_{\mathrm{i}}/\eta_{\mathrm{t}})\,\lambda_{\mathrm{i}}\,,$$

i.e. the wavelength is smaller in media with a higher refractive index. Notice that according to this equation, the product $\lambda\eta = \lambda_0$ is constant across the interface, where λ_0 is the wavelength of light in a vacuum.

For spectral radiance with respect to wavelength, Kirchhoff's equilibrium radiance law now states that the quantity

$$\frac{L_\lambda(\mathbf{x}, \omega, \lambda_0/\eta)}{\eta(\mathbf{x}, \omega, \lambda_0)^3}$$

is constant throughout the enclosure. This equation applies separately at each wavelength λ_0. The factor of η^3 instead of η^2 occurs because L_λ is defined as a derivative with respect to wavelength (see [Nicodemus 1976, p. 51]). Effectively, when light enters a medium of higher refractive index, the same light energy is squeezed into a smaller band of wavelengths, which causes the spectral radiance to increase proportionately.

Applying this version of Kirchhoff's equilibrium radiance law, the reciprocity condition for BSDF's becomes

$$\frac{f_{\mathrm{s},\lambda}(\omega_{\mathrm{i}} \!\to\! \omega_{\mathrm{o}}, \lambda_0/\eta_{\mathrm{i}})}{\eta_{\mathrm{o}}(\lambda_0)^3} \;=\; \frac{f_{\mathrm{s},\lambda}(\omega_{\mathrm{o}} \!\to\! \omega_{\mathrm{i}}, \lambda_0/\eta_{\mathrm{o}})}{\eta_{\mathrm{i}}(\lambda_0)^3}\,, \tag{6.4}$$

where the wavelength parameter of $f_{\mathrm{s},\lambda}(\omega_{\mathrm{i}} \!\to\! \omega_{\mathrm{o}}, \lambda)$ refers to the incident light, and λ_0 is the wavelength in a vacuum.

6.3 Conservation of energy

We show that the following condition is implied by conservation of energy:

Theorem 6.2. *If f_s is the BSDF for a physically valid surface, which is either the boundary of an opaque object or the interface between two non-absorbing media, then*

$$\int_{\mathcal{S}^2} f_\mathrm{s}(\omega_\mathrm{i} \to \omega_\mathrm{o})\, d\sigma^\perp(\omega_\mathrm{o}) \;\le\; 1 \qquad \textit{for all} \quad \omega_\mathrm{i} \in \mathcal{S}^2\,. \tag{6.5}$$

This is very similar to the energy-conservation condition for BRDF's, which was mentioned in Section 3.6.2.

Proof. Equation (6.5) can be proven from the relations

$$\begin{aligned} E &= \int_{\mathcal{S}^2} L_\mathrm{i}(\omega)\, d\sigma^\perp(\omega) \\ L_\mathrm{o}(\omega_\mathrm{o}) &= \int_{\mathcal{S}^2} L_\mathrm{i}(\omega)\, f_\mathrm{s}(\omega \to \omega_\mathrm{o})\, d\sigma^\perp(\omega) \\ M &= \int_{\mathcal{S}^2} L_\mathrm{o}(\omega)\, d\sigma^\perp(\omega)\,, \end{aligned}$$

where E denotes the irradiance (i.e. the incident power per unit area), and M denotes the radiant exitance (i.e. the scattered power per unit area, see Section 3.4). By conservation of energy, we require that $M \le E$ for all possible incident radiance functions L_i; that is, the surface should never scatter more light than it receives.

To obtain the desired condition (6.5), we fix a particular direction ω_i, and consider the incident radiance distribution $L_\mathrm{i}(\omega) = \delta_{\sigma^\perp}(\omega - \omega_\mathrm{i})$, i.e. we let the incident power be concentrated in a single direction ω_i.[5] With this choice of L_i, we obtain

$$\begin{aligned} E &= \textstyle\int_{\mathcal{S}^2} L_\mathrm{i}(\omega)\, d\sigma^\perp(\omega) &&= 1 \\ L_\mathrm{o}(\omega_\mathrm{o}) &= \textstyle\int_{\mathcal{S}^2} L_\mathrm{i}(\omega)\, f_\mathrm{s}(\omega \to \omega_\mathrm{o})\, d\sigma^\perp(\omega) &&= f_\mathrm{s}(\omega_\mathrm{i} \to \omega_\mathrm{o}) \\ M &= \textstyle\int_{\mathcal{S}^2} L_\mathrm{o}(\omega)\, d\sigma^\perp(\omega) &&= \textstyle\int_{\mathcal{S}^2} f_\mathrm{s}(\omega_\mathrm{i} \to \omega_\mathrm{o})\, d\sigma^\perp(\omega_\mathrm{o})\,, \end{aligned}$$

from which the requirement that $M \le E$ gives the desired result. ∎

[5] Alternatively, we could use a sequence of radiance functions that approximate L_i, to avoid the issue of whether L_i is allowed to be a Dirac distribution.

Appendix 6.A Helmholtz reciprocity

In the graphics and radiometry literature, the symmetry of BRDF's is often attributed to the Helmholtz reciprocity principle. Apparently this notion first arose in the radiometry literature (see the references in Section 6.A.3), and migrated to graphics through the work of Nicodemus (e.g. see [Nicodemus et al. 1977, p. 40], [Nicodemus 1965, p. 769]).

In this section, we examine the original statement of Helmholtz reciprocity, and show that it does not imply the symmetry of BRDF's. Helmholtz stated his principle only for classical optical systems (consisting of mirrors and lenses), and thus with regard to the reflection of light from surfaces, his principle applies only to mirrors. He does not mention non-specular reflection of any sort (e.g. diffuse or glossy surfaces). (Of course, we would not expect Helmholtz to mention BRDF's in any case, since the concept of a BRDF was not invented at that time.)

6.A.1 Summary of the principle

The Helmholtz reciprocity principle is found in his famous treatise on physiological optics, first published in 1856 [von Helmholtz 1856, p. 231]. This three-volume work concerns human vision: the anatomy of the eye, the mechanisms of sensation, and the interpretation of those sensations. With regard to optics, Helmholtz' main concern was to analyze the properties of the eye within the framework of classical geometric optics.

In this context, Helmholtz proposed the following reciprocity principle for beams traveling through an optical system (i.e. a collection of mirrors, lenses, prisms, etc). Suppose that a beam of light A undergoes any number of reflections or refractions, eventually giving rise (among others) to a beam B whose power is a fraction f of beam A. Then on reversing the path of the light, an incident ray B' will give rise (among others) to a beam A' whose power is the same fraction f of beam B'.[6] In other words, the path of a light beam is always reversible, and furthermore the relative power loss is the same for propagation in both directions.

The main point is that the only type of reflection considered by Helmholtz is specular reflection from mirrors. Thus, his principle does not have any direct implications for general BRDF's (or BSDF's).

Note that Helmholtz reciprocity can easily be extended to materials that are composed of many

[6] Our paraphrasing follows that of Chandrasekhar [1960, p. 176].

small mirrors. By considering the limit as these mirrors become very small, a variety of interesting materials can be obtained (this is the basic idea behind *microfacet reflection models* [Torrance & Sparrow 1967, Cook & Torrance 1982]). However, note that this approach is not adequate to prove a reciprocity principle for real materials, since it only applies to a particular *model* for reflection from surfaces. That is, real surfaces are not necessarily composed of microfacets, so this type of argument cannot be used to make statements about the properties of real BRDF's. (Microfacet models were only proposed as a model to explain reflection from metals, and in any case the microfacets are generally so small that geometric optics is not applicable: diffraction theory must be used instead [He et al. 1991].)

6.A.2 Helmholtz' original statement

We now examine the original statement of Helmholtz reciprocity[7], to explain in more detail why it applies only to specular reflection. Note that the following quotation is for polarized light, which makes it slightly more complicated. With respect to the paraphrasing of the previous section, it assumes that beam A is polarized in a given plane P_A, and that we only measure the component of beam B that is polarized in a given plane P_B. Helmholtz' principle then states that the same fraction of power is lost for a beam traveling in either direction:

> Suppose light proceeds by any path whatever from a point A to another point B, undergoing any number of reflections or refractions *en route*. Consider a pair of rectangular planes a_1 and a_2 whose line of intersection is along the initial path of the ray at A; and another pair of rectangular planes b_1 and b_2 intersecting along the path of the ray when it comes to B. The components of the vibrations of the aether particles in these two pairs of planes may be imagined. Now suppose that a certain amount of light J leaving the point A in the given direction is polarised in the plane a_1, and that of this light the amount K arrives at the point B polarised in the plane b_1; then it can be proved that, when the light returns over the same path, and the quantity of light J polarised in the plane b_1 proceeds from the point B, the amount of this light that arrives at the point A polarised in the plane a_1 will be equal to K.
>
> Apparently the above proposition is true no matter what happens to the light in the way of single or double refraction, reflection, absorption, ordinary dispersion, and diffraction, provided that there is no change of its refrangibility, and provided it does not traverse any magnetic medium that affects the position of the plane of polarisation, as Faraday found to be the case.

[7] Translated from the German [von Helmholtz 1856, p.231]. A somewhat shorter statement of this principle appears in [von Helmholtz 1903, Section 42, p. 158], but the apparent meaning is the same.

It is absolutely clear that Helmholtz did not intend this principle to be applied to diffuse reflection. In the cited reference, he consistently uses the word *reflection* to mean only specular (mirror-like) reflection. For example, in the discussion leading up to the statement of the principle above [von Helmholtz 1856, p. 230], Helmholtz states and proves a similar theorem where it is obvious that only specular reflection is considered. The very phrase *reflections and refractions* implies specularity, since otherwise the theorem would be stated in terms of reflection and *transmission*.

Also, the principle refers to the *amounts* of light leaving A and arriving at B. In the terminology of the day, an "amount" of light referred to total flux or power.[8] Thus, the principle does not even make sense for diffuse reflection: the power arriving on beam B would always be zero, since only an infinitesimal quantity of power is reflected by a diffuse surface in any particular direction. In order to make sense for non-specular reflection, the law would need to relate the power at A to a *differential* quantity at B, such as irradiance. (This is exactly what was done by Lord Rayleigh, in the reciprocity principle discussed below.)

Another important fact is that Helmholtz reciprocity is not always valid, as we will discuss in Appendix 6.D). Interestingly, Helmholtz did not provide a proof of his principle, claiming that "anybody who is at all familiar with the laws of optics can easily prove it for himself" [von Helmholtz 1856, p. 231].

6.A.3 Further reading on Helmholtz reciprocity

This section gives a sampling of the various sources of information available concerning reciprocity principles. It is by no means exhaustive.

First, there are references that interpret the Helmholtz reciprocity principle correctly, in the limited sense discussed above (beams propagating through an optical system, rather than general scattering from surfaces). These include Planck [1914, p. 49], von Fragstein [1955], Chandrasekhar [1960, p. 176], and Born & Wolf [1986, p. 381].

Second, there are sources in the radiometry literature that claim (in passing) that Helmholtz reciprocity implies the symmetry of physically valid BRDF's. These include McNicholas [1928], de Vos [1954], de la Perrelle et al. [1963], Nicodemus [1965, p. 769], and Nicodemus et al. [1977, p. 40].

[8] It is important that Helmholtz stated his law in terms of power, rather than radiance, since this way his law is valid even when A and B lie in media with different refractive indices. If it were stated in terms of radiance (which Helmholtz calls "brightness"), there would need to be an η^2 scaling factor as discussed in Section 5.2. Helmholtz was aware of this scaling factor [von Helmholtz 1856, p. 233], and thus phrased his law to make it as general as possible.

Another worthwhile reference is Minnaert [1941], who recognized that the original statement of Helmholtz reciprocity does not apply to general scattering, but shows how the statement can be reinterpreted to have greater generality.

Third, there are general reciprocity principles in the physics literature regarding the scattering of electromagnetic waves. These include Kerr [1987], Saxon [1955], and de Hoop [1960]. These principles are unrelated to Helmholtz', and require additional physical assumptions for their validity.

This brings up an important point, which is that many reciprocity principles in optics are derived by *starting* with a relationship of the sort that we want to prove (i.e. an assumption that is equivalent to the symmetry of BRDF's). For example, several reciprocity principles for volume scattering are proven in [Case 1957], by assuming that the phase function is symmetric (bottom p. 653). By making additional assumptions of this sort (e.g. that all scattering particles have random orientations), it is possible to derive a wide variety of reciprocity principles in optics [van de Hulst 1957, Chapter 5], [Hovenier 1969], [van de Hulst 1980, Chapter 3]. It would be an easy mistake to derive a reciprocity principle for BSDF's by starting with results such as these.

Appendix 6.B Lord Rayleigh's reciprocity principle

In 1900, Lord Rayleigh stated a reciprocity principle for non-specular reflection, and apparently he was the first to do so.[9] Essentially, this principle asserts that real materials have symmetric BRDF's. The original statement is as follows:

> Suppose that in any direction (i) and at any distance r from a small surface (S) reflecting in *any manner* there be situated a radiant point (A) of given intensity, and consider the intensity of the reflected vibrations at any point B situated in direction ϵ and at distance r' from S. The theorem is to the effect that the intensity is the same as it would be at A if the radiant point were transferred to B. [*Footnote:* I have not thought it necessary to enter into questions connected with polarization, but a more particular statement could easily be made.]

Translated into modern terminology, we are given a small reflective surface, exposed to a small light source and a small *irradiance sensor* (which measures the power per unit area falling on a square facing toward the reflective surface). His principle states that if the positions of the source and sensor are exchanged, the measured irradiance will be the same. This implies that the corresponding BRDF must be symmetric, as may easily be verified.

Observe that Rayleigh's principle is merely a statement of fact; no proof was given in terms of more basic physical laws. Although it was claimed as a consequence of "a fundamental principle of reciprocity, of such generality that escape from it is difficult" (to be found in his *Theory of Sound* [Rayleigh 1877, Sec. 109, p. 154]), the methods used there are not rigorous by modern standards, and are not explicitly related to light. Furthermore, they require symmetry assumptions about the underlying system (e.g. see [Rayleigh 1877, Sec. 103a, p. 139]) that seem no more justifiable than assuming the symmetry of the BRDF in the first place.

[9] This observation was made by Chandrasekhar [Chandrasekhar 1960, p. 177]. Rayleigh's statement of reciprocity can be found in a short letter to the *Philosophical Magazine* [Rayleigh 1900, p. 324] (reprinted in [Rayleigh 1964, p. 480]).

Appendix 6.C Time reversal invariance and the irreversibility of light scattering

The symmetry of BRDF's is sometimes attributed to a physical law known as *time reversal invariance*. In this section, we explain why such claims are incorrect. Time reversal invariance does not have any direct consequences for the symmetry of BRDF's, because most observable light scattering processes are *irreversible*.

We first explain the principle of time reversal invariance, which applies the laws of physics at a microscopic level (e.g. interactions between individual particles). Next, we explain why observable processes are usually *irreversible*, even though they are governed by microscopically reversible laws. Finally, we explain how this applies to light scattering: we show that observable light scattering processes are almost always irreversible, so that time reversal invariance does not have any direct implications for the symmetry of BRDF's.

Time reversal. It is known that the fundamental laws of physics are invariant under the operation of *time reversal*, in which the time variable is negated in formulas and equations. More precisely, this principle should be called *motion reversal invariance*, since it asserts that if the motions of all particles and waves in a system are reversed, then they will retrace their former paths [de Groot & Mazur 1962, p. 35]. This principle holds in any physical system, as long as there are no external magnetic fields; otherwise, the direction of the field must be reversed along with the wave and particle motions, in order for time reversal invariance to hold ([de Groot & Mazur 1962, p. 38], [de Groot 1963]).[10]

This principle can be stated more precisely as follows. Let A and B be any two *microscopic states* of the given system, where each state completely specifies the attributes of all particles and waves. We let $p(A, B, t)$ denote the *transition function* for this system, i.e. the probability density that if the system is in state A, it will evolve to state B over a time interval of length t. (Note that according to quantum mechanics, the universe is not deterministic; thus, we can only compute the *probability* with which the system evolves from state to state.) Finally, given a state X, we let $-X$ denote the state obtained by motion reversal of all particles and waves (including the reversal of magnetic fields, if necessary). Given these definitions, the principle of time reversal invariance then

[10]Technically, there are some known exceptions to time reversal invariance, however these involve nuclear interactions and are not significant for optics [*Brittanica Online* 1996].

states that

$$p(A, B, t) \;=\; p(-B, -A, t) \qquad \text{for all } A,\, B, \text{ and } t$$

(see [de Groot 1963]).

Irreversible processes. It is important to realize that time reversibility applies only at a microscopic level, and that most physically observable processes are *irreversible*. As a simple example of an irreversible process, consider a box that is divided into two compartments, one containing vacuum and another filled with air. If a hole is made in the separating wall, air will rush from one side to the other until the pressures on both sides are the same. This process is irreversible, since if the motions of all the particles are reversed, they will not revert to their original configuration.

How can we explain this paradox, given that the underlying physical laws are time reversible? Briefly, the reason is that observable states and microscopic states are not in one-to-one correspondence. In fact, each observable state X' can be realized in a large number of different microscopic ways, all of which are indistinguishable with respect to measurable properties (such as pressure or temperature). This idea is closely related to the concept of *entropy*: letting W denote the number of ways that an observable state X' can be realized, its entropy is given by $S = k \ln W$, where k is the Boltzmann constant. Thus, states with higher entropy can be realized in a greater number of microscopic ways.

Given these facts, irreversible processes can arise as follows. Consider a discrete system where there are only 100 microscopic states $X_1, \ldots, X_{100}$, and the transition probabilities between them are all equal: $p(X_i, X_j) = p(X_j, X_i)$ for all i and j. We suppose that motion reversal is simply the identity operation, i.e. $-X_i = X_i$ (recalling that $-X$ denotes motion reversal). Clearly this system is microscopically time reversible, since

$$p(A, B) \;=\; p(-B, -A) \qquad \text{for all } A \text{ and } B\,.$$

However, now suppose that the system has only two observable states A' and B', which correspond to 1 and 99 microscopic states respectively. It is easy to verify that $p(A', B') = 99/100$, while $p(B', A') = 1/100$. Thus from an observable point of view, the system is not time reversible: if the system moves from A' to B', and motion reversal is applied to the microscopic state underlying B', then the system is far more likely to move to another microscopic state of B', than it is to return to the original microscopic state underlying A'. The general reason for this behavior is that B' corresponds to a much larger number of microscopic states than A': that is, it has a higher entropy.

In a real thermodynamic system, any measurable increase in entropy corresponds to a huge increase in the number of equivalent microscopic states. For all practical purposes, the probability of returning to the original observable state is zero, and thus any process which increases entropy is said to be *irreversible*process. This is the essence of the second law of thermodynamics: given a closed, isolated system, it will always move from less probable to more probable observable states. (These are only the basic ideas behind irreversible processes; for more rigorous arguments, see van Kampen [1954], de Groot & Mazur [1962], or de Groot [1963].)

Irreversibility of light scattering. As we mentioned, the symmetry of BRDF's is sometimes attributed directly to the time reversibility of physical laws. This is incorrect, because the scattering of light at an ordinary surface is irreversible. (Here we are referring to the *observable* behavior of light scattering, which corresponds to an average behavior over many indistinguishable microscopic states.) There are two reasons for this: first, when a light beam strikes a surface, some of the energy is absorbed (and converted into heat). Motion reversal of all photons and other particles will not convert this heat back into light. Second, the incident beam will generally be scattered in many directions (e.g. by a diffuse surface); and if the direction of this scattered light is reversed to form an incident distribution, it does not recreate the original beam. Both of these situations represent an increase in entropy, and are not reversible.[11] Thus, time reversal invariance does not have any direct implications for ordinary BRDF's, where some light is absorbed and/or scattered in multiple directions.

Light scattering is only reversible at a perfect mirror (if such a thing could be constructed), or at an optically smooth interface between two dielectric materials, as pointed out by Stokes in 1849 (see [Lekner 1987, p. 36] or [Knittl 1962]). It is occasionally claimed that Maxwell's equations themselves are time reversible, but this is true only in special cases. Obviously Maxwell's equations are not time reversible in general, since they describe phenomena such as absorption.[12]

Although time reversal invariance is not useful to us directly, recall that it underlies the principle of detailed balance. Since detailed balance holds even for irreversible processes, it can be applied to light scattering, as we did in Section 6.2. The main limitation of detailed balance (as compared to time reversal) is that it only holds for systems in thermodynamic equilibrium.

[11] See [Jones 1953] for an intuitive discussion of the irreversibility of light scattering; however, note that this paper has a few technical errors.

[12] Although Maxwell's equations are not invariant under the operation of time reversal, they do have other symmetry properties. This has been studied by Šantavý [1961], who describes an operation closely related to time reversal under which Maxwell's equations are indeed invariant.

Appendix 6.D Exceptions to reciprocity

When we derived the reciprocity condition (6.3) for BSDF's, we needed two important assumptions: that there are no external magnetic fields, and that there are no absorbing media.[13] In this section, we give some insight into why these assumptions are necessary, by showing what goes wrong when they are violated.

First, we discuss magnetic fields, which cause problems for polarized light. Then we discuss absorbing media, which at first appear to violate not only the reciprocity condition (6.3), but also the principles of detailed balance and conservation of energy. We explain the apparent contradictions, and we also derive a reciprocity condition that applies to absorbing media (in Section 6.D.2.3).

6.D.1 Magnetic fields and the Faraday effect

We show how magnetic fields can cause reciprocity principles to fail. This includes both Helmholtz reciprocity, and the reciprocity condition (6.3) for general BSDF's.

The source of these problems is the *Faraday effect*, which states that when plane-polarized light propagates within an external magnetic field, the plane of polarization is rotated. For example, consider a polarized beam of light that passes through an electromagnet. According to the Faraday effect, the plane of polarization will rotate in the same direction as the current flow in the magnet. This rotation does not depend on the direction of light propagation, but only on the magnetic field: thus, if the same beam is reflected back and forth through the magnet, the rotation increases each time. This obviously represents an exception to the Helmholtz reciprocity principle, as it was stated for polarized light, and Helmholtz himself was aware of this (see the quotation in Section 6.A.2).

Lord Rayleigh's light trap. As a more dramatic example of how reciprocity can fail, Lord Rayleigh proposed the following *light trap*. Consider a horizontal cylinder filled with a magnetic medium,[14] together with an external field such that polarized light passing through the cylinder is rotated by 45 degrees. Now suppose that a polarizer is placed at either end of the cylinder, oriented so that their planes of polarization are 45 degrees apart. In this situation, light passing in one direction through the cylinder is completely blocked by the second polarizer, while light traveling the

[13] An *absorbing medium* is one that absorbs some of the light energy passing through it, so that the intensity of a light beam decreases with distance.

[14] Note that the Faraday effect only occurs in substances that are *magnetically active*. Oxygen, hydrogen, and water are all magnetically active to some degree [Born & Wolf 1986, p. 3], although the Faraday effect is strongest in substances such as carbon bisulphide [Drude 1900, Chapter 7].

other direction is transmitted. Thus, a source of light at one end of the cylinder A would be visible at the other end B, while a source at B would not be visible from A.

Using this idea, we can show the existence of BSDF's that do not obey the reciprocity condition (6.3) in the presence of a magnetic field. For example, consider a thin film of magnetized iron. When plane-polarized light passes through this film, its plane of polarization is rotated slightly [Drude 1900, p. 451]. Now suppose that the iron is coated with polarizing films on both sides; this will produce an effect similar to the "light trap", i.e. the transmissivity of the surface will be different for light traveling through it in opposite directions. The same can be achieved for reflective surfaces, for example by coating a magnetized mirror with certain kinds of optical crystals [Drude 1900].

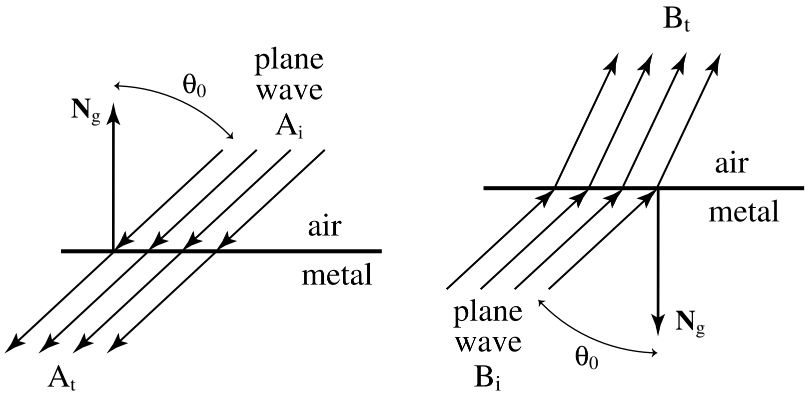
6.D.2 Transmission between absorbing media

Reciprocity also fails when light is transmitted into *absorbing media*. For these media, the radiance of a light beam decreases exponentially with the distance traveled. The absorption is due to electrical conduction, which transforms light energy into electron vibrations (which then appear as heat). The medium may be only slightly absorbing, as with imperfect dielectric materials, or it may be a conductor (metals), in which case light is virtually extinguished after propagating only a few wavelengths.

For absorbing media, there are two separate ways in which reciprocity fails [von Fragstein 1955]. We give a brief introduction to them here, and provide more detail in the following sections.

First, the path of a light beam is not always reversible. For example, consider a light wave that is transmitted from air into metal (Figure 6.2). For some metals, there exists an non-zero angle of incidence where the transmitted beam does not change its direction (i.e. it is not refracted), and yet light beams that go in the opposite direction from metal into air are refracted for all non-zero incident angles. For other metals, the reverse is true: beams are always refracted upon entering the metal, but for beams exiting the metal, there is a non-zero angle where the direction of propagation does not change. Note that these situations do not happen in the familiar case of non-absorbing media, where light beams are refracted for all non-zero angles of incidence, and the path of a light beam is always reversible.

The second effect concerns the *transmissivity* of the interface between two media, i.e. the fraction of incident power that is transmitted through the surface. For absorbing media, a larger fraction of light can be transmitted in one direction than the other. Letting $\tau_{i,j}$ denote the transmissivity from



(a) Air to metal **(b)** Metal to air

Figure 6.2: When absorbing media such as metals are present, the path of a light beam is not always reversible. For example, when a light beam $A_{\rm i}$ is transmitted from air into some metals, there is a non-zero angle of incidence θ_0 for which the beam does not change its direction of propagation (Figure (a)). However, a beam of light B_i traveling in the reverse direction (from metal into air) is refracted at the surface, and follows a different path (Figure (b)).

medium i to medium j, the transmissivities in opposite directions are related by

$$\frac{\tau_{1,2}}{\tau_{2,1}} = \frac{1+\kappa_1^2}{1+\kappa_2^2},$$

where κ_i is the *attenuation index* for medium i (which measures the rate of light absorption in that medium).[15] Furthermore, the reflectivity and transmissivity at such an interface can sum to more than one (which cannot happen with non-absorbing media).

At first sight, these properties appear to violate the principles of detailed balance and conservation of energy, respectively. However, this is not the case. In the following sections, we explain these apparent contradictions, and we also derive a more general reciprocity condition for BSDF's that holds even when there are absorbing media.

[15] The attenuation index is defined so that when a light wave travels a single wavelength λ, its amplitude is reduced by a factor of $e^{-2\pi\kappa}$. This is not the same as the *absorption coefficient* $\sigma_{\rm a}$ used in the volume rendering and radiation transport literature, which measures the rate of absorption per unit length. The two quantities are related by $\sigma_{\rm a} = 4\pi\kappa/\lambda$ [Born & Wolf 1986, p. 614]. Adding further to the confusion, κ is often called the *extinction coefficient*, which is the same name given in the transport literature to the sum of the absorption and scattering coefficients.

6.D.2.1 Non-reversibility of optical paths

We have stated that the path of a light beam between air and metal is not always reversible. To explain this, consider a homogeneous plane wave A_i traveling through the air (i.e. an infinitely wide beam, propagating in a given direction). Suppose that this wave strikes the planar boundary of an absorbing medium, where the angle of incidence is $\theta(A_\mathrm{i}) = \theta_0$ (see Figure 6.2(a)).

In [von Fragstein 1955], it is shown that for some metals, there is a non-zero value of θ_0 for which the incoming wave will not be refracted (i.e. $\theta(A_\mathrm{t}) = \theta(A_\mathrm{i}) = \theta_0$). On the other hand, a homogeneous wave B_i traveling in the reverse direction *will* be refracted; that is, if it strikes the boundary at an angle $\theta(B_\mathrm{i}) = \theta_0$ from inside the metal, it will exit at an angle $\theta(B_\mathrm{t})$ that is different from θ_0 (Figure 6.2(b)).

At first sight, this appears to contradict the principle of detailed balance. At thermodynamic equilibrium, we must have the same energy flowing both ways between any given pair of directions; thus, it would seem that if the wave A_i is not refracted, then the wave B_i should not be refracted as well. (Otherwise, power arriving from the given direction ω_i would be scattered to $-\omega_\mathrm{i}$, but not vice versa.)

The crucial observation is that the two situations we have considered are actually *not* the reverse of each other. To obtain the refraction results above, the waves A_i and B_i must both be *homogeneous*, i.e. their amplitude must be constant along each wavefront [von Fragstein 1955]. However, when the wave A_i is refracted into metal, the result A_t is not a homogeneous wave: the wavefronts are perpendicular to the direction of propagation $\theta(A_\mathrm{t}) = \theta_0$, while the surfaces of constant amplitude are parallel to the boundary between the two media [Born & Wolf 1986, p. 616]. This happens because each point on a given wavefront has traveled a different distance through the absorbing medium, and the amplitude of the wave falls off according to the distance traveled.

Because of this, the irreversibility of optical paths between absorbing media is a bit misleading. The situation considered by von Fragstein is not a true reversal of the optical path, because he assumes that the incident wave is homogeneous in both directions. Suppose that instead, we let B_i be an inhomogeneous wave of the same type as A_t, where the wavefronts are perpendicular to the direction θ_0, but the surfaces of constant amplitude are parallel to the boundary. It is possible to show that this yields a transmitted wave B_t of the same form as A_i: a homogeneous wave, propagating in the desired direction θ_0. Thus, the requirements of detailed balance are satisfied.

To see that this is true, consider the following experiment. Suppose that the metal forms a thin layer, with air on both sides, and consider a homogeneous wave A_i that is incident at the angle θ_0. This wave enters the metal, where it is refracted into an inhomogeneous wave A_t traveling in the same direction. This wave propagates to the far side of the metal layer, where we will rename it B_i, and is then transmitted into air yielding a wave $B_{\rm t}$. This wave $B_{\rm t}$ must clearly be homogeneous, since all parts of the wave have traversed the same thickness of metal. Furthermore, from Snell's law it is straightforward to show that the waves $A_{\rm i}$ and $B_{\rm t}$ have the same direction of propagation [Born & Wolf 1986, p. 629]. Thus, the second refraction (which acts on an inhomogeneous wave $B_{\rm i}$) exactly reverses the action of the first, as we claimed above.

6.D.2.2 Apparent non-conservation of energy

Next, we turn to the transmissivity of the interface between absorbing media. We have claimed that the reflectivity and transmissivity at such an interface can sum to more than one, which appears to violate conservation of energy. This can be explained in terms of interference between the incident and reflected light waves.

In particular, consider a planar boundary between air and metal, where the metal has an attenuation index of κ. Suppose that an incident wave $A_{\rm i}$ strikes the boundary from within the metal, giving rise to a reflected wave $A_{\rm r}$ and a transmitted wave $A_{\rm t}$. Then according to von Fragstein [1950], the transmissivity satisfies

$$\tau \;=\; (1+\kappa^2)(1-\rho)\,,$$

where the reflectivity ρ can be determined from the Fresnel laws [Born & Wolf 1986, p. 628]. Note that the factor $1+\kappa^2$ can be rather large; e.g. for silver it is approximately 400 [von Fragstein 1950, p. 65]. Thus, the amount of transmitted light can be much larger than it would be if $\rho+\tau=1$.

To understand this, we must examine the definitions of reflectivity and transmissivity. They measure the power of the reflected and transmitted waves, as compared to the incident wave:

$$\rho \;=\; \frac{\Phi(A_{\rm r})}{\Phi(A_{\rm i})}\,, \qquad \tau \;=\; \frac{\Phi(A_{\rm t})}{\Phi(A_{\rm i})}\,.$$

However, the key observation is that the incident and reflected waves are propagating in the same medium, and that these two waves can *interfere* with each other. The power carried toward the boundary by the combined wave $A_{\rm i}+A_{\rm r}$ can be either more or less than the intuitively expected value $\Phi(A_{\rm i})-\Phi(A_{\rm r})$.

For transmission from metal to air, the combined wave $A_{\rm i}+A_{\rm r}$ carries more power toward the boundary than expected. This can be shown from Maxwell's equations, where the additional energy appears as *mixed product* terms in the Poynting vector [von Fragstein 1950]. In the one-dimensional case [Salzberg 1948], the power carried by the combined wave can be written in the form

$$(1/2)\,\mathrm{Re}(E_{\mathrm{i}} H_{\mathrm{i}}^*) - (1/2)\,\mathrm{Re}(E_{\mathrm{r}} H_{\mathrm{r}}^*) - (1/2)\,\mathrm{Re}(E_{\mathrm{i}} H_{\mathrm{r}}^* - E_{\mathrm{r}} H_{\mathrm{i}}^*)\,,$$

where E and H denote the complex amplitudes of the electric and magnetic components of the wave, respectively, and * denotes complex conjugation. The first two terms denote the power of the incident and reflected waves, while the third term measures the additional energy flow due to interference.

Effectively, the interference between incident and reflected waves causes there to be less absorption in the metal near the boundary [von Fragstein 1955]. That is, for a wave propagating far inside the metal, absorption will occur at the usual rate (as determined by κ). However, as the wave approaches the boundary, the rate of absorption becomes smaller, due to interference from the reflected wave. Thus, when the wave finally exits from the metal, it will have much more power than it would if the reduced absorption were not taken into account.

6.D.2.3 A reciprocity condition for BSDF's with absorbing media

We derive a reciprocity condition for BSDF's that applies even when absorbing media are present. This requires only one small change to the argument in Section 6.2.

Recall that for a system in thermodynamic equilibrium, where only non-absorbing media are present, that the quantity L/η^2 is constant throughout the enclosure. When absorbing media are present, this must be modified: it is possible to show that the quantity

$$\frac{L(\mathbf{x}, \omega)\,(1 + \kappa^2)}{\eta^2}$$

is constant throughout the enclosure [von Fragstein 1950, Tingwaldt 1952], where L, κ, and η are parameterized by frequency ν. Thus according to this formula, the equilibrium radiance is smaller in an absorbing medium than in a non-absorbing one.

By repeating the argument of Section 6.2, we can now show that an arbitrary, physically valid BSDF must satisfy

$$\frac{f_{\mathrm{s}}(\omega_{\mathrm{i}} \to \omega_{\mathrm{o}})\,(1 + \kappa_{\mathrm{o}}^2)}{\eta_{\mathrm{o}}^2} = \frac{f_{\mathrm{s}}(\omega_{\mathrm{o}} \to \omega_{\mathrm{i}})\,(1 + \kappa_{\mathrm{i}}^2)}{\eta_{\mathrm{i}}^2}\,, \tag{6.6}$$

where all quantities are parameterized by frequency ν. This is clearly a generalization of the condition (6.3) given for non-absorbing media.

To put this into perspective, however, the difference between (6.6) and (6.3) is utterly insignificant for the typical participating media used in graphics, because the attenuation indices are so small.

For example, consider a medium that is so dense that 99% of the incident light is absorbed after a distance of one millimeter. This corresponds to an attenuation index of only $\kappa = 0.00037$ (for light with a wavelength of 500nm), so that the $(1 + \kappa^2)$ change in transmissivity is inappreciable.

Also, note that the participating media in graphics are often not true absorbing media, but instead consist of small particles (e.g. clouds, fog, smoke). These materials are described by *scattering theory* [van de Hulst 1957], rather than the theory of absorbing media described here. (In a true absorbing medium, the particles must be of negligible size compared to the wavelength: for example, an iodine solution, or a cloud of chlorine gas.)

6.D.2.4 Discussion

Given these bizarre examples, it is clear that absorbing media cannot be described with familiar optical concepts. The idea of independent waves propagating and reflecting, each with its own power, is simply meaningless in an absorbing medium [Salzberg 1948]. For example, consider the standard Fresnel formulas for reflection and transmission between absorbing media [Born & Wolf 1986, p. 628]. According to these formulas, there is non-zero reflection even at a ficticious interface between two *identical* media; furthermore, the corresponding transmissivity is greater than one. To handle such situations correctly, it is necessary to work with explicit wave descriptions (e.g. monochromatic waves described by their phase and amplitude), rather than with secondary concepts such as power.

It is reassuring to note that the strange effects we have described are restricted to the absorbing media themselves. For example, consider a wave A_{i} that is transmitted from air, through a metal film, and then back into air to yield a wave A_{t}. It can be shown that the optical path is reversible, and that the transmissivity of the film is the same in both directions [Lekner 1987]. This holds even if the film consists of many layers of absorbing and non-absorbing media (known as a *striated medium*).