Microfacet Based Bidirectional Reflectance Distribution Function

Morten S. Mikkelsen
Naughty Dog Inc., USA

November 2, 2009

Abstract

In computer graphics, analytical physically based specular BRDFs are most often expressed by a surface distribution function and a Fresnel term. In some cases a geometric attenuation factor is included. The first such model to be introduced from optics, to computer graphics, was the Torrance–Sparrow model which represents a distribution of small mirror–like facets. In this paper we analyze the three–term model by deriving it from an explicit microfacet formulation. Furthermore, it is common to combine such a specular BRDF with a diffuse term to account for multiple reflections and subsurface scattering. We discuss how this can be done without violating physical properties such as energy conservation and symmetry.

Contents

1 Introduction

As light travels it interacts with the medium in which it is moving. These interactions are referred to as *scattering* and *absorption* where the former causes the light to stray from its path and move in a new direction and the latter is when the light is absorbed by the medium. For a more thorough description the reader is referred to [Ish78].

Certain media, such as air, are very sparse which keeps the quantity of interactions per unit volume low. Other media such as wood or metal are very dense in comparison and thus most interactions occur close to the surface. At the heart of computer graphics we need to be able to determine the fraction of reflected outgoing light from a contribution of incoming light. The redistribution of light is specific to properties of the material and so to produce realistic images it is necessary to know the characteristics of the redistribution function associated with the material. This is known as the *bidirectional reflectance distribution function* (**BRDF**) and was first defined by Fred Nicodemus in 1965 and is described in [NRH$^+$77]. The function takes two directions as input $\vec{\omega}_i$ and $\vec{\omega}_o$ which are both given relative to the orientation of the surface defined by $\vec{n}$. The returned value is the ratio of differential reflected *radiance* exiting in the direction $\vec{\omega}_o$ to the differential *irradiance* incident on the surface from direction $\vec{\omega}_i$.

Though the terminology and framework of the BRDF is given in [NRH$^+$77] it is in practice not trivial to determine a compact description of the function such that it captures real world materials and remains efficient to sample. For this reason a wide range of more or less approximate solutions have been introduced. Some aim to reproduce real–world materials accurately by performing measurements made using the gonioreflectometer [GTS$^+$97]. Others with the same goal have used image–based techniques to produce BRDF measurements [War92], [MWLT00] and [MPBM03].

The problem for such a data driven approach to the BRDF is both size of data and the complexity in sampling it. For this reason other authors have focused on finding an analytical function which is reminiscent of some BRDFs. An example is the phenomenological, yet widely used for its simplicity, BRDF associated with the Phong lighting model [Pho75]. Another example is a physically–based model known as the Torrance–Sparrow [TS67]. Their expression is derived based on the premise that the surface of real world objects is variably rough and irregular. This model was introduced to the computer graphics community by Jim Blinn in 1977 [Bli77] and has been the source of inspiration and subsequent research for many years. There are many examples but to name a few the interested reader is referred to papers such as [Bli77], [CT81], [War92], [Sch94] and [AP07].

The Torrance–Sparrow BRDF model is represented, primarily, by three terms. The first is a Fresnel term which is well understood and is discussed briefly in section 2.1. The second is known as a geometric attenuation factor which requires no user–defined parameters. Finally, the third is a probability distribution function (PDF) which describes the distribution of orientations, of

the normal, across the projected area of some infinitesimal patch surface. This patch is interpreted as a dense set of connected microfacets at which level reflectance is determined by the Fresnel term. This concept is discussed in greater detail in section 2.3. The reason light from one direction is, generally, not reflected into a single outgoing direction is because of the microfacets which have different orientations. For a normally distributed surface a rough patch yields a wide base. On the other hand a patch which is almost planar has a thin PDF.

Though the Gaussian function is used as a PDF in [TS67] the BRDF model is flexible enough to allow any choice of a hemispherical PDF. To summarize, the Torrance–Sparrow model is popular because it is: physically based, relatively simple and analytical. Furthermore, it is flexible in regards to choice of PDF.

As one might expect much of the subsequent research has been on which PDF to choose and also how to find the best fit relative to some form of input data [CT81], [War92] and [AP07]. In this paper we will not focus on choice of PDF or fitting. This is an analysis on the three–term model and it is shown how it is derived, in section 2.4, from an explicit microfacet formulation which is given in section 2.3. The goal is to gain an overall insight into the structure of the model. In section 2.5 we provide a clarification and a derivation of the geometric attenuation factor. It is common practice to add a diffuse term to the BRDF to account for multiple reflections and local subsurface scattering. In section 2.6 we discuss a way to do this without violating physical properties (given in section 2.2) of the general BRDF.

2 Theory

The Torrance–Sparrow BRDF model [TS67] consists, primarily, of a Fresnel term, a visibility term, and a hemispherical PDF. This analytical model assumes that the surface consists of small mirror–like facets. These will be referred to as microfacets. An exact evaluation of outgoing radiance resulting from incident light on a microfacet based surface involves integration and has a different formulation. Several mathematical steps are involved before the approximate form is reached and in this section we will provide a derivation.

Initially, we introduce in section 2.1 the basic radiometric terminology which is used in this paper. In section 2.2 we discuss the definition of the general BRDF and the physical properties which it must obey. From this, definition, the exact formulation of the microfacet based BRDF is given in section 2.3 where it initially is assumed light is reflected in a single bounce. Subsequently, we show in section 2.4 how to derive the Torrance–Sparrow model from the exact form. In order to achieve this an important approximation is made where it is assumed that interception of incoming and outgoing light is determined by a local model. This is a significant step since visibility is known to be a global problem. The issue is discussed in section 2.5 and a derivation of the local model is given. It is common in both optics and computer graphics to use a diffuse term to account for multiple reflections and subsurface scattering. We discuss in section 2.6 a way to do this without violating the necessary physical

properties.

2.1 Radiance

In radiometry, the power Φ of light is known as radiant flux and is measured in Watts. To know the radiant flux reflected by a visible surface we need to relate this quantity of light to parameters such as: the direction $\vec{\omega}$ from which the surface receives light, the orientation of the surface, which is given by its normal $\vec{n}$ and the area dA from which the light is reflected. Finally, also the differential frequency of the light though this dependency will, for brevity, remain implied in our equations. The concept is known as *radiance*, $L(x, \vec{\omega})$, and the relation between this and power is given by 7–1 in [Ish78] as

$$\frac{d^2\Phi}{dAd\vec{\omega}} = (\vec{\omega} \bullet \vec{n})L(x, \vec{\omega}) \tag{1}$$

Here the symbol $\bullet$ denotes the dot product which gives us scalar projection of the incoming light from $\vec{\omega}$ and onto the surface orientation $\vec{n}$. The density of power is diminished as these orientations are misaligned.

There exists an ambiguity in regards to radiance. It can either move along the path of $\vec{\omega}$ or it can be coming from $\vec{\omega}$. In this paper radiance received from the direction of $\vec{\omega}$ is called *incoming radiance* $L(x, \vec{\omega})$ and radiance moving along $\vec{\omega}$ is *outgoing radiance* $L_o(x, \vec{\omega})$. Consequently, it follows that

$$L_o(x, -\vec{\omega}) = L(x, \vec{\omega}) \tag{2}$$

An important element in computer graphics (and general optics) is *reflectance* which is the fraction of incident power that is reflected at the surface of a medium. For a perfectly flat surface, of a thick homogeneous medium, reflectance is described by the *Fresnel* equations which were deduced by Augustin-Jean Fresnel. The Fresnel equations are derived from the Maxwell equations based on ideal circumstances where both media are homogeneous, i.e., constant material properties and the *interface* between them is planar which gives a mirror–like reflection. We will in the following let $F_r\left(\vec{n} \bullet \vec{\omega}\right)$ denote the Fresnel reflection associated with unpolarized incident light coming from ω and the transmittance is given by $F_t\left(\vec{n} \bullet \vec{\omega}\right) = 1 - F_r\left(\vec{n} \bullet \vec{\omega}\right)$. A thorough description of the Fresnel equations is beyond the scope of this paper so the interested reader is referred to pp. 240–250 in [Fl28]. Furthermore, we will let the dependency on the index of refraction $\mu(\lambda) \in \mathbb{C}$ remain implicit to achieve a more compact notation.

In practice most materials do not exhibit a mirror–like reflection but tend to distribute received light into multiple outgoing directions. The issue is discussed in the following section.

2.2 Bidirectional Reflectance Distribution Function

The form of reflection described in the previous section applies to ideal circumstances where the surface is entirely planar and exhibits no subsurface scattering.

In practice most objects have at a microscale level a rough and irregular high frequency profile. This is the case even for objects which appear to be planar. These surface irregularities cause the incident radiance to reflect into multiple outgoing directions. Furthermore, non-metallic materials tend to exhibit some degree of *subsurface scattering*. This implies that light transmitted into a medium may, after some number of scattering events, return to the surface and transmit back out. This behavior will have an additional impact on the resulting directional distribution of outgoing radiance. Additionally, the quantity of light absorbed by the medium is wavelength dependent. Thus the portion of transmitted flux which resurfaces, as outgoing radiance, depicts coloring by the medium.

To model the redistribution of flux, at the surface, of a medium which exhibits subsurface scattering one must take into account that the entry and exit positions are not the same which makes the problem inherently more difficult to solve. The interested reader is referred to papers such as [JMLH01] and [DJ05] which model subsurface scattering for piecewise homogeneous media. In our work we will restrict our analysis to surface scattering specifically. The relation between received and reflected light is described by the *bidirectional reflectance distribution function* (BRDF) $f_r(x, \vec{\omega}_o, \vec{\omega}_i)$, introduced by [NRH$^+$77], which multiplied by the incident differential flux gives the corresponding differential outgoing radiance. The position x is two–dimensional since it is a position at the surface of the medium which means f_r is a six–dimensional function. The BRDF is often studied at a specific surface point or simply assumed to be position-invariant. Thus it reduces to a four–dimensional function and we will proceed under this assumption.

Given a known radiance field $L(x, \vec{\omega}_i)$, since f_r relates radiance and flux, the outgoing radiance is given by the following integral

$$L_o(x, \vec{\omega}_o) = \int_{\Omega_{2\pi}} f_r(\vec{\omega}_o, \vec{\omega}_i) L(x, \vec{\omega}_i)(\vec{\omega}_i \bullet \vec{n}) d\vec{\omega}_i \tag{3}$$

since $L(x, \vec{\omega}_i)(\vec{\omega}_i \bullet \vec{n})$ is the incident flux given by equation (1).

The BRDF has two important properties which are the Helmholtz reciprocity and the second is energy conservation (see [NRH$^+$77]). The former tells us that the throughput, of light, is the same going from $\vec{\omega}_i$ into $\vec{\omega}_o$ as it is coming from $\vec{\omega}_o$ and into $\vec{\omega}_i$. Intuitively, this is because the throughput of light relates to the measure of surface area which is mutually visible, across dA, and has a normal which approximately aligns with the vector half–way between ω_i and $\vec{\omega}_o$. For multiple bounces a similar concept applies since the path is mutual. This property of symmetry gives us $f_r(\vec{\omega}_o, \vec{\omega}_i) = f_r(\vec{\omega}_i, \vec{\omega}_o)$.

The second property which is energy conservation tells us that

$$\int_{\Omega_{2\pi}} f_r(\vec{\omega}_i, \vec{\omega}_o)(\vec{\omega}_o \bullet \vec{n}) d\vec{\omega}_o \leq 1 \tag{4}$$

This can be explained by considering the incident flux received from within $d\vec{\omega}_i$ as specified by some arbitrary intensity

$$k = L(x, \vec{\omega}_i)(\vec{\omega}_i \bullet \vec{n})$$

Thus using the BRDF and equation (1) the ratio between differential flux reflected and the differential incident flux is

$$\begin{aligned}\frac{\frac{d^2\Phi_o}{dAd\vec{\omega}_o}}{\frac{d^2\Phi_i}{dAd\vec{\omega}_i}} &= \frac{(\vec{\omega}_o \bullet \vec{n}) f_r(\vec{\omega}_i, \vec{\omega}_o) k}{k} \\ &= (\vec{\omega}_o \bullet \vec{n}) f_r(\vec{\omega}_i, \vec{\omega}_o)\end{aligned}$$

and as we see k is canceled. By integrating over all possible outgoing directions we get the ratio between outgoing flux and incident flux from $\vec{\omega}_i$. This is identical to the left side of inequality (4) which must be less than one since the material reflects, due to surface scattering, no more than that which is received.

If we differentiate equation (3) on both sides, with respect to $\vec{\omega}_i$, and rearrange terms we obtain the following expression for the BRDF

$$f_r(\vec{\omega}_o, \vec{\omega}_i) = \frac{1}{L(x, \vec{\omega}_i)(\vec{\omega}_i \bullet \vec{n})} \frac{dL_o(x, \vec{\omega}_o)}{d\vec{\omega}_i} \tag{5}$$

which deserves some clarification. The last term represents outgoing radiance resulting from flux received within a unit solid angle around $\vec{\omega}_i$. Alternatively, we may think of $d\vec{\omega}_i$ as an infinitesimal solid angle and $dL_o(x, \vec{\omega}_o)$ as the outgoing radiance that is obtained from flux received within $d\vec{\omega}_i$. Thus in equation (5) we have, in the denominator, flux per unit area received from within $d\vec{\omega}_i$ and in the numerator we have $dL_o(x, \vec{\omega}_o)$.

In this section we have discussed the BRDF and the properties required for a possible solution. In the following section we will determine an expression for (5) by assuming that the Fresnel equations are used to determine reflectance at a microscale level.

2.3 Explicit Microfacet–Based BRDF Formulation

As mentioned in section 2.2 real world materials tend to reflect incident light into multiple outgoing directions. In geometrical optics [TS67] this is explained by the presence of microfacets and reflectance is determined using the Fresnel equations at each facet. A differential surface S thus consists of some dense tesselation into such facets. The mean surface dA is referred to as the macro surface and is the projected area of the micro surface S. As the incoming radiance strikes S, reflectance will occur, into multiple outgoing directions since S is not planar. Let the reflection of $\vec{\omega}$ by some normal $\vec{n}$ be given by

$$R_{\vec{n}}(\vec{\omega}) = 2(\vec{n} \bullet \vec{\omega})\vec{n} - \vec{\omega} \tag{6}$$

If a beam of light (specified by the radiance $L(x_j, \omega_i)$) is incident on a planar infinitesimal surface (such as a microfacet) of area dx_j and normal $\vec{n}_j$. The power received, by the facet, from within some small solid angle $d\vec{\omega}_i$ is determined by equation (1) as

$$\Phi = (\vec{n}_j \bullet \vec{\omega}_i) L(x_j, \omega_i) d\vec{\omega}_i dx_j$$

The notation x_j, where $j \in \mathbb{N}$, is used as an indication that though we use continuous notation we imagine some dense, yet finite, tesselation of S into microfacets.

It was explained at the end of section 2.1 that the portion of received light, which is reflected into the direction $R_{\vec{n}_j}(\vec{\omega}_i)$, is determined by the factor $F_r(\vec{n}_j \bullet \vec{\omega}_i)$. Thus we can accumulate the power reflected, in a single bounce, across all of the surface area S using the following

$$\Phi_{ref} = \int_S F_r(\vec{n}_j \bullet \vec{\omega}_i)\, V(x_j, \vec{\omega}_i, R_{\vec{n}_j}(\vec{\omega}_i))\, (\vec{n}_j \bullet \vec{\omega}_i)\, dx_j \cdot L(x, \vec{\omega}_i) d\vec{\omega}_i \qquad (7)$$

where we have used that $L(x_j, \vec{\omega}_i) \simeq L(x, \vec{\omega}_i)$ since the extent of S is infinitesimal. However, a microfacet dx_j may be occluded by S in the incident direction $\vec{\omega}_i$ or the reflected flux in the outgoing direction may be intercepted. For this reason we have the mutual visibility term $V(x_j, \vec{\omega}_i, R_{\vec{n}_j}(\vec{\omega}_i))$ during integration.

Now our goal is to express the BRDF, in equation (5), given our differential surface S of microfacets. To do this we must determine an expression for $dL_o(x, \vec{\omega}_o)$ which is in the numerator of equation (5). The differential operator refers to an implicit dependence on the differential solid angle of incident light $d\vec{\omega}_i$. Equation (1) applies to both the incoming and outgoing radiance. Though the projected differential areas dA_i and dA_o are strictly not the same we are determining $dL_o(x, \vec{\omega}_o)$ from $L(x, \vec{\omega}_i)$ specifically. Thus since we are not dealing with subsurface scattering we can assume that $dA_i = dA_o$ and we will refer to this as simply dA. Conceptually this is the flat mean region, through S, centered at x and with normal $\vec{n}$. Furthermore, dA is the parallel projected area of the micro surface S.

Now to determine $dL_o(x, \vec{\omega}_o)$ the radiant flux reflected in equation (7) is interpreted as having been reflected by the macro surface dA, with normal $\vec{n}$, which corresponds to S. Furthermore, since we are dealing with radiance we must filter out reflected flux which is not within the solid angle $d\vec{\omega}_o$. We do this by scaling, during integration, by the term $\delta\left(R_{\vec{n}_j}(\vec{\omega}_i) \in d\vec{\omega}_o\right)$. In a sense we should be asking to what extent the reflection of the incoming solid angle $d\vec{\omega}_i$ is within the outgoing solid angle $d\vec{\omega}_o$ but since the integral in equation (3) is over $\vec{\omega}_i$ the corresponding solid angle is infinitesimal. Thus by the principal of integration only the **direction** $\vec{\omega}_i$ is queried. We find $dL_o(x, \vec{\omega}_o)$ using equations (1) and (7) which gives

$$dL_o(x, \vec{\omega}_o) = \frac{\int_S F_r(\vec{n}_j \bullet \vec{\omega}_i)\, \delta\left(R_{\vec{n}_j}(\vec{\omega}_i) \in d\vec{\omega}_o\right) (\vec{n}_j \bullet \vec{\omega}_i)\, V dx_j \cdot L(x, \vec{\omega}_i) d\vec{\omega}_i}{(\vec{n} \bullet \vec{\omega}_o)\, d\vec{\omega}_o dA} \qquad (8)$$

where $V = V(x_j, \vec{\omega}_i, \vec{\omega}_o)$. Finally, we evaluate the BRDF by substituting equation (8) into equation (5) which gives us the resulting equation

$$f_r(\vec{\omega}_o, \vec{\omega}_i) = \frac{\int_S F_r(\vec{n}_j \bullet \vec{\omega}_i)\, \delta\left(R_{\vec{n}_j}(\omega_i) \in d\vec{\omega}_o\right) (\vec{n}_j \bullet \vec{\omega}_i)\, V dx_j}{(\vec{n} \bullet \vec{\omega}_i)\, (\vec{n} \bullet \vec{\omega}_o)\, d\vec{\omega}_o dA} \qquad (9)$$

As one might expect the factors $L(x, \vec{\omega}_i)$ and $d\vec{\omega}_i$ are cancelled since the BRDF

describes the redistribution of flux regardless of the incoming density. However, it remains an assumption that $d\vec{\omega}_i$ is an infinitesimal solid angle.

By the rule of reciprocity the BRDF must obey $f_r(\vec{\omega}_o, \vec{\omega}_i) = f_r(\vec{\omega}_i, \vec{\omega}_o)$ and showing that this applies to equation (9) is possible through a few key observations. Though the outgoing solid angle is conceptually arbitrarily small we can consider the size of it the same whether the incident direction of radiance is from $\vec{\omega}_i$ or $\vec{\omega}_o$. Thus given equation (6) we have $R_{\vec{n}_j}(\vec{\omega}_i) \bullet \vec{\omega}_o = R_{\vec{n}_j}(\vec{\omega}_o) \bullet \vec{\omega}_i$ which gives $\delta\left(R_{\vec{n}_j}(\omega_i) \in d\vec{\omega}_o\right) = \delta\left(R_{\vec{n}_j}(\omega_o) \in d\vec{\omega}_i\right)$. Let the vector half–way between the incident and the outgoing direction be denoted

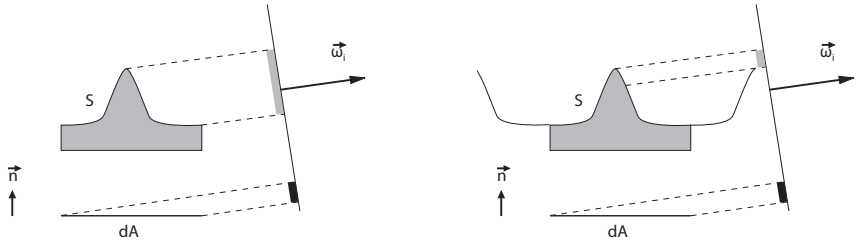
$$\vec{h} = \frac{\vec{\omega}_i + \vec{\omega}_o}{\|\vec{\omega}_i + \vec{\omega}_o\|} \tag{10}$$

Next since the outgoing solid angle is small we have that $\vec{n}_j \simeq \vec{h}$ which in return gives $\vec{n}_j \bullet \vec{\omega}_i \simeq \vec{n}_j \bullet \vec{\omega}_o$ which then gives $F_r(\vec{n}_j \bullet \vec{\omega}_i) \simeq F_r(\vec{n}_j \bullet \vec{\omega}_o)$. This completes the symmetry of equation (9).

As previously mentioned an additional requirement is that energy conservation is obeyed which is defined by inequality (4). To show this we insert the microfacet based BRDF, given by equation (9), into the left side of inequality (4).

$$\begin{aligned}
\int_{\Omega_{2\pi}} f_r(\vec{\omega}_o, \vec{\omega}_i)(\vec{\omega}_o \bullet \vec{n}) d\vec{\omega}_o &= \frac{\int_S F_r(\vec{n}_j \bullet \vec{\omega}_i)(\vec{n}_j \bullet \vec{\omega}_i) V(x_j, \vec{\omega}_i, \vec{\omega}_o) dx_j}{(\vec{n} \bullet \vec{\omega}_i) dA} \\
&\leq \frac{\int_S (\vec{n}_j \bullet \vec{\omega}_i) V(x_j, \vec{\omega}_i) dx_j}{(\vec{n} \bullet \vec{\omega}_i) dA} \qquad (11) \\
&\leq 1 \qquad (12)
\end{aligned}$$

In inequality (11) we have used that $F_r(\vec{n}_j \bullet \vec{\omega}_i) \leq 1$ and that $V(x_j, \vec{\omega}_i, \vec{\omega}_o) \leq V(x_j, \vec{\omega}_i)$. The last inequality (12) is based on a few observations which will be clarified in the following. The projected area of the micro surface S is denoted dA and the mean normal is the unit vector $\vec{n}$. Thus the denominator in inequality (11) represents scalar projection onto the plane with normal $\vec{\omega}_i$ as shown at the bottom of figure 1(a). The same projection of all microfacets dx_j on S is given by the numerator. Overlaps due to wrinkles, on S, are accounted for by the visibility term $V(x_j, \vec{\omega}_i)$. The surface of the grey medium in figure 1(a) represents S and dA is shown at the bottom. The projection of S is illustrated by a solid grey rectangular region. A similar block in black is shown for the projection of the macro surface dA. We would like the coverage of S to be less than that of dA such that inequality (12) is true. This is clearly not the case in figure 1(a). The reason is that we must also take a previous assumption into account which is that the BRDF is position independent. This and the observation that S is a differential surface indicates that S repeats across the material. The distinction is shown in figure 1(b) where the projection of the visible part of the medium is reduced due to occlusion. Given the interpretation that S repeats the maximum visible range is from peak to peak. This range is equal to the horizontal extent of S which is equal to dA. Thus the numerator is less than or equal to the denominator which means energy conservation is obeyed.



(a) single surface

(b) repeated surface

Figure 1: The surface of the grey medium represents the differential patch S and is shown in figure 1(a). It is parallel projected onto the view–plane as seen from the direction $\vec{\omega}_i$. The coverage is shown in the figure. If the surface is repeated visibility of the surface of the medium is reduced as shown in figure 1(b). The line dA represents the projected extent of S.

It will be discussed in section 2.4 how the current microfacet based BRDF, given by equation (9), can be simplified into an analytical expression without integration.

2.4 Distribution–Based BRDF Formulation

In numerous papers [TS67], [Bli77] and [CT81], the microfacet based BRDF is presented as an analytical expression comprising three primary terms: The Fresnel factor $F_r(\vec{\omega}_o \bullet \vec{h})$, a geometry term G and finally a chosen surface distribution function D which must obey the following

These papers do not derive this expression, but present it, in this form based on intuitive arguments. For better clarity we will show, in this section, how it can be derived from the explicit microfacet distribution based BRDF given by equation (9).

The Torrance–Sparrow model [TS67] is, perhaps, the most famous example of the form since this was the first to be introduced to the computer graphics community and for its flexibility regarding choice of surface distribution function D.

$$f_r(\vec{\omega}_o, \vec{\omega}_i) \simeq \frac{D(\vec{h})G(\vec{\omega}_i, \vec{\omega}_o)F_r(\vec{\omega}_o \bullet \vec{h})}{4\,(\vec{n} \bullet \vec{\omega}_i)\,(\vec{n} \bullet \vec{\omega}_o)} \tag{13}$$

This is the modern day interpretation, in computer graphics, of what the Torrance-Sparrow BRDF looks like (see equation 9.8 in [PH04]) and interestingly this exact equation never appears in the original paper. The reason for this is, possibly, that it was aimed at the optics community and not computer graphics. Regardless, when we refer to the Torrance–Sparrow model we are referring to

equation (13). Note that a valid surface distribution function must obey the following

$$\int_{\Omega_{2\pi}} D(\vec{h})(\vec{h} \bullet \vec{n}) d\vec{h} = 1 \tag{14}$$

and describes the distribution of surface orientations.

In equation (9) the term $\delta\left(R_{\vec{n}_j}(\omega_i) \in d\vec{\omega}_o\right)$ ensures that contributions of flux are only received when the surface normal $\vec{n}_j$ reflects the incident direction $\vec{\omega}_i$ into the outgoing solid angle $d\vec{\omega}_o$. The set of surface normals which obey this are contained within a solid angle $d\vec{h}$ around the half–way vector defined by equation (10). There exists a relation between these two solid angles and it is given by

$$d\vec{\omega}_o = 4 \cdot (\vec{h} \bullet \vec{\omega}_o) \cdot d\vec{h} \tag{15}$$

This relation is also given in the Torrance–Sparrow paper [TS67] and for a proof the interested reader is referred to appendix D of [NIK92].

We will now proceed with showing how the Torrance–Sparrow (13) can be derived from equation (9). Our first step is to substitute equation (15) into (9) and next to perform integration over the macro surface. We do this by performing integration by subsitution using $da_j = (\vec{n}_j \bullet \vec{n})\, dx_j$ which gives

$$\begin{aligned} f_r(\vec{\omega}_o, \vec{\omega}_i) &= \frac{\int_S F_r(\vec{n}_j \bullet \vec{\omega}_i)\, \delta\left(\vec{n}_j \in d\vec{h}\right)(\vec{n}_j \bullet \vec{\omega}_i)\, V dx_j}{4(\vec{h} \bullet \vec{\omega}_o)(\vec{n} \bullet \vec{\omega}_i)(\vec{n} \bullet \vec{\omega}_o)\, d\vec{h} dA} \\ &= \frac{\int_{dA} F_r(\vec{n}_j \bullet \vec{\omega}_i)\, \delta\left(\vec{n}_j \in d\vec{h}\right) \frac{\vec{n}_j \bullet \vec{\omega}_i}{\vec{n}_j \bullet \vec{n}} V da_j}{4(\vec{h} \bullet \vec{\omega}_o)(\vec{n} \bullet \vec{\omega}_i)(\vec{n} \bullet \vec{\omega}_o)\, d\vec{h} dA} \end{aligned} \tag{16}$$

For the first step we used that $R_{\vec{n}_j}(\omega_i) \in d\vec{\omega}_o \Leftrightarrow \vec{n}_j \in d\vec{h}$ and for the second step there is an assumption that every point on dA maps to only one point on S. Thus the set of normals which exist on S represent some subset $U \subseteq \Omega_{2\pi}$. Let a specific map be given, $\varphi : dA \mapsto U$, which for any point on dA produces the surface normal of the corresponding point on S. Let the area measure of a *well behaved* subset of dA be denoted μ. We know that any subset of $M \subseteq U$ has the property that the preimage is contained in $\varphi^{-1}(M) \subseteq dA$. We will assume that the preimage is well behaved, when M is, which means φ is a measurable function. Thus according to measure and integration theory $\mu\left(\varphi^{-1}(M)\right)$ is a measure on U. We can extend this measure to all of $\Omega_{2\pi}$ by the observation that the preimage of the complement is $\varphi(\Omega_{2\pi} \backslash U) = \varnothing$. Thus let the chosen measure be zero for any subset contained in the complement.

Every term in the integrand, in equation (16), is a function of the surface normal except for $V = V(x_j, \vec{\omega}_i, \vec{\omega}_o)$. This is because visibility by nature is a global problem and not a local one. In order to apply the substitution $\vec{n}_j = \varphi(p_j)$, where $p_j \in dA$, we need to use an approximation of V where visibility is determined based on local properties such as the surface normal $\vec{n}_j$ and the incident and outgoing direction $\vec{\omega}_i$ and $\vec{\omega}_o$.

$$V'(\vec{n}_j, \vec{\omega}_i, \vec{\omega}_o) \simeq V(x_j, \vec{\omega}_i, \vec{\omega}_o)$$

Such an approximation is possible by the assumption that S represents a dented flat surface. These dents are referred to as cavities and these intuitively allow for a local estimate. We will discuss the explicit formulation of V' in section 2.5 but for now assume it is known. We can now rewrite equation (16) into the following

$$\begin{aligned} f_r(\vec{\omega}_o, \vec{\omega}_i) &\simeq \frac{\int_{\Omega_{2\pi}} F_r\left(\vec{n}_j, \vec{\omega}_i\right) \delta\left(\vec{n}_j \in d\vec{h}\right) \frac{\vec{n}_j \bullet \vec{\omega}_i}{\vec{n}_j \bullet \vec{n}} V' \mu\left(\varphi^{-1}(d\vec{n}_j)\right)}{4(\vec{h} \bullet \vec{\omega}_o)\left(\vec{n} \bullet \vec{\omega}_i\right)\left(\vec{n} \bullet \vec{\omega}_o\right) d\vec{h} dA} \\ &= \frac{\int_{d\vec{h}} F_r\left(\vec{n}_j, \vec{\omega}_i\right) \frac{\vec{n}_j \bullet \vec{\omega}_i}{\vec{n}_j \bullet \vec{n}} V' \mu\left(\varphi^{-1}(d\vec{n}_j)\right)}{4(\vec{h} \bullet \vec{\omega}_o)\left(\vec{n} \bullet \vec{\omega}_i\right)\left(\vec{n} \bullet \vec{\omega}_o\right) d\vec{h} dA} \end{aligned}$$

such that integration is now performed over $d\vec{h}$. The probability of finding a normal in $d\vec{n}_j$ relates directly to the area measure of the preimage. Subsequently, an applied PDF should approximate the following

$$p(\vec{n}) = \frac{\mu\left(\varphi^{-1}(d\vec{n})\right)}{dA d\vec{n}} \tag{17}$$

Note that this reference equation is normalized for integration over $\Omega_{2\pi}$ and that the product $p(\vec{n})d\vec{n}$ describes the probability that the surface normal is in $d\vec{n}$. By inserting the reference equation we finally arrive at

$$\begin{aligned} f_r(\vec{\omega}_o, \vec{\omega}_i) &\simeq \frac{dA \int_{d\vec{h}} F_r\left(\vec{n}_j, \vec{\omega}_i\right) \frac{\vec{n}_j \bullet \vec{\omega}_i}{\vec{n}_j \bullet \vec{n}} V' p(\vec{n}_j) d\vec{n}_j}{4(\vec{h} \bullet \vec{\omega}_o)\left(\vec{n} \bullet \vec{\omega}_i\right)\left(\vec{n} \bullet \vec{\omega}_o\right) d\vec{h} dA} \\ &= \frac{F_r(\vec{\omega}_o \bullet \vec{h}) \cdot V'(\vec{h}, \vec{\omega}_i, \vec{\omega}_o) \cdot p(\vec{h})}{4\left(\vec{n} \bullet \vec{\omega}_i\right)\left(\vec{n} \bullet \vec{\omega}_o\right)\left(\vec{n} \bullet \vec{h}\right)} \end{aligned} \tag{18}$$

where we have used, in the second step, that integration is over the differential solid angle $d\vec{h}$ which we also have in the denominator. This provides the form in equation (13) since the first parameter given to V' is $\vec{h}$ which is expressed by equation (10). Thus we may consider V' a function of $\vec{\omega}_i$ and $\vec{\omega}_o$ only as is the case for the term G in equation (13). The term D is equivalent to $\frac{p(\vec{h})}{\vec{n} \bullet \vec{h}}$ as indicated by equation (14) and because p is normalized over $\Omega_{2\pi}$.

A common example of a surface distribution function D is the normalized Phong distribution where n is the distribution parameter.

$$D(\vec{h}) = \frac{n+2}{2\pi}(\vec{h} \bullet \vec{n})^n \tag{19}$$

Another example is the Beckmann surface distribution with distribution parameter m.

$$\begin{aligned} D(\vec{h}) &= \frac{e^{-\frac{\tan^2\left(\arccos(\vec{h} \bullet \vec{n})\right)}{m^2}}}{\pi m^2 (\vec{h} \bullet \vec{n})^4} \\ &= \frac{e^{-\frac{1-(\vec{h} \bullet \vec{n})^2}{(\vec{h} \bullet \vec{n})^2 m^2}}}{\pi m^2 (\vec{h} \bullet \vec{n})^4} \end{aligned} \tag{20}$$

Both equations (19)–(20) obey equation (14). This implies the π factor is conceptually misplaced in the Cook–Torrance model [CT81] since it belongs in the surface distribution function (20). Furthermore, the Cook–Torance model is equivalent to the Torrance–Sparrow model (13), using Bechmann, but scaled by a factor of four which is most likely an error.

The Beckmann distribution is more costly to evaluate than the normalized Phong model. For this reason it is worth pointing out that the parameter m, in Beckmann, can be converted to n in the normalized Phong model by equating the normalization constants of the two distributions (19) and (20) which gives

$$n(m) = \frac{2}{m^2} - 2 \tag{21}$$

This provides a very close approximation from a visual standpoint when $n > 8$.

In summary, to arrive at the form (18) the approximation was used that there exists a local model to determine mutual visibility. We will discuss the approximation used by the Torrance–Sparrow model in the following section.

2.5 Approximation of the Visibility Factor

It was mentioned in section 2.4 how the product $p(\vec{h})d\vec{h}$ represents the fraction of dA which has a surface normal in $d\vec{h}$. Ideally, multiplying by the visibility function V', should give us the fraction of this which is mutually visible. That is the following equation

$$V'(\vec{\omega}_i, \vec{\omega}_o) = \frac{\int_{\varphi^{-1}(d\vec{h})} V(x_j, \vec{\omega}_i, \vec{\omega}_o) da_j}{\mu\left(\varphi^{-1}(d\vec{h})\right)} \tag{22}$$

It should be mentioned that the preimage $\varphi^{-1}(d\vec{h})$ is generally a disconnected set.

As previously discussed visibility is a global problem so to make it a local problem the Torrance–Sparrow model assumes the surface S represents a dented plane such that each dent is given by a V shaped groove with the sides at equal but opposite angles to the mean surface normal. Since the V–groove must belong to the preimage of $d\vec{h}$ we can assume that one of the opposing sides has its normal within $d\vec{h}$. This side is the reflector of the incident light and the other side will be referred to as the opposing side. If light is intercepted by the opposing side, as shown in figure 2(a), it is said to have been masked. If the light is intercepted before it reaches the reflector it is said to have been shadowed. We can see this by imagining the flux path to be reversed in figure 2(a). Simultaneous masking and shadowing is possible but as pointed out in [TS67] one or the other will dominate.

If we disregard multiple V–grooves within the preimage, for now, then given our reference equation (22) the goal is to determine the fraction of the reflecting side which reflects light unintercepted. This is shown in figure 2(a) where a V–groove is shown from the side. The two sides are both of length l and that

which is masked has length m. Thus as specified in [TS67] the result is $1-\left(\frac{m}{l}\right)$. However, expressing V' in terms of m and l is impractical so in [Bli77] it is shown how the same result can be expressed using the directions $\vec{n}$, $\vec{h}$, $\vec{\omega}_i$ and $\vec{\omega}_o$ as input and we will show this in the following but using a more intuitive proof.

Let the projection of $\vec{\omega}_o$ into the plane spanned by $\vec{n}$ and $\vec{h}$ be denoted $\vec{\omega}'_o$. This is the direction in which the reflected light exits and is shown in figure 2(b). By applying the law of sines to the triangle depicted in the figure we have

$$\frac{m}{l} = \frac{\sin(b)}{\sin(a)} \tag{23}$$

And next we can express this as cosines by using $sin(v) = \cos\left(\frac{\pi}{2} - v\right)$ which used in equation (23) gives

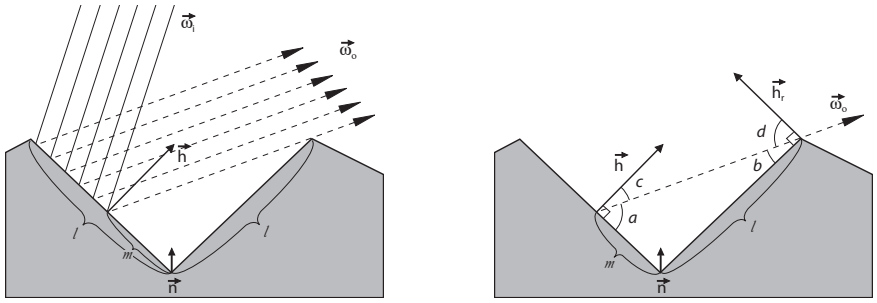
$$\begin{aligned} \frac{m}{l} &= \frac{\cos(d)}{\cos(c)} \\ &= \frac{-\vec{\omega}'_o \bullet \vec{h}_r}{\vec{\omega}'_o \bullet \vec{h}} \end{aligned} \tag{24}$$

The vector $\vec{h}_r$ is the normal of the opposing side which is

$$\vec{h}_r = R_{\vec{n}}\left(\vec{h}\right)$$

the reflection of $\vec{h}$ by the mean normal $\vec{n}$. Using equation (6) we now get

$$1 - \frac{m}{l} = \frac{2\left(\vec{n} \bullet \vec{h}\right)\left(\vec{\omega}'_o \bullet \vec{n}\right)}{\vec{\omega}'_o \bullet \vec{h}}$$



(a) ratio (b) geometrical constellation

Figure 2: The percentage of unintercepted reflected light in figure 2(a) is $1 - \frac{m}{l}$. Angles are given in figure 2(b).

As we see this result does not depend on the magnitude or sign of $\vec{\omega}_o'$. Let the unit normal to the plane spanned by the vectors $\vec{n}$ and $\vec{h}$ be given by the vector

$$\vec{v} = \frac{\vec{n} \times \vec{h}}{\|\vec{n} \times \vec{h}\|}$$

as we mentioned initially $\vec{\omega}_o'$ is the projection of $\vec{\omega}_o$ into this plane and since we can ignore the magnitude this projection is given by

$$\vec{\omega}_o' = \vec{\omega}_o - \vec{v} \cdot (\vec{v} \bullet \vec{\omega}_o)$$

Thus since $\vec{v}$ is perpendicular to both $\vec{n}$ and $\vec{h}$ we finally get

$$1 - \frac{m}{l} = \frac{2\left(\vec{n} \bullet \vec{h}\right)(\vec{\omega}_o \bullet \vec{n})}{\vec{\omega}_o \bullet \vec{h}}$$

which is identical to the result in [Bli77]. The equation gives us the result when masking is the dominant form of interception. A similar scenario when shadowing dominates can be seen by imagining the flux path is reversed in figure 2(a) and as pointed out in [Bli77] we can use the same equation to determine the result but with the roles of $\vec{\omega}_o$ and $\vec{\omega}_i$ exchanged. Since $\vec{\omega}_i \bullet \vec{h} = \vec{\omega}_o \bullet \vec{h}$ we arrive at the final equation

$$V'(\vec{\omega}_i, \vec{\omega}_o) = \min\left(1, \frac{2\left(\vec{n} \bullet \vec{h}\right)(\vec{\omega}_o \bullet \vec{n})}{\vec{\omega}_o \bullet \vec{h}}, \frac{2\left(\vec{n} \bullet \vec{h}\right)(\vec{\omega}_i \bullet \vec{n})}{\vec{\omega}_o \bullet \vec{h}}\right) \tag{25}$$

As we see this result is independent of the size of the V–groove and its location on the plane. Thus the result is the same for an arbitrary amount of V–grooves where one side has its normal in $d\vec{h}$. Furthermore, the result is conveniently determined from dot products between vectors which are known during the time of rendering.

In these past sections we have discussed the redistribution of light which reflects in a single bounce off of a surface of microfacets. In section 2.6 we will discuss how to deal with light which is transmitted or is subjected to multiple reflections.

2.6 Multiple Reflections

It is comprehensive to determine the redistribution of light due to multiple reflections and/or internal scattering of transmitted light. For this reason the Torrance–Sparrow paper [TS67] chooses to model both by a diffuse component which represents a uniform distribution. It seems unlikely that second order reflections would yield a uniform distribution but according to [TS67] this model is in general accord with measurements made by several authors. Since it is a very practical approximation it is widely used today in the computer graphics community.

Let $f_{r,s}$ represent light reflected in a single bounce and let the diffuse component be represented by

$$f_{r,d}(\vec{\omega}_i, \vec{\omega}_o) = \frac{1}{\pi} \tag{26}$$

Note that this definition obeys energy conservation since

$$\begin{aligned} \int_{\Omega_{2\pi}} f_{r,d}(\vec{\omega}_o, \vec{\omega}_i)(\vec{\omega}_o \bullet \vec{n}) d\vec{\omega}_o &= \frac{1}{\pi} \int_{\Omega_{2\pi}} (\vec{\omega}_o \bullet \vec{n}) d\vec{\omega}_o \\ &= 1 \end{aligned}$$

The full BRDF f_r is obtained by accumulating both contributions. Furthermore, let two user–defined constants be given $k_s, k_d \in [0; 1]$. The typical BRDF, in computer graphics, is thus given by

$$f_r(\vec{\omega}_i, \vec{\omega}_o) = k_s f_{r,s}(\vec{\omega}_i, \vec{\omega}_o) + k_d f_{r,d}(\vec{\omega}_i, \vec{\omega}_o)$$

A detail which is often ignored or forgotten is that this composition should still obey energy conservation. A sensible remedy may be obtained by the observation that only light which has not reflected in a single bounce should be redistributed by $f_{r,d}$. We can achieve this by introducing a function which determines the diffuse reflectance given some incident direction $\vec{\omega}$

$$\rho(\vec{\omega}) = \int_{\Omega_{2\pi}} f_{r,s}(\vec{\omega}', \vec{\omega})(\vec{\omega}' \bullet \vec{n}) d\vec{\omega}' \tag{27}$$

This was explained in section 2.2 as the percentage of flux, received from a specific direction, which is reflected (into any direction). Thus the remainder, which is not reflected in a single bounce, is given by $1 - k_s\rho(\vec{\omega}_i)$ and we can scale this onto the diffuse component to make the BRDF energy preserving. However, doing so violates symmetry. To preserve both properties we arrive at the following expression

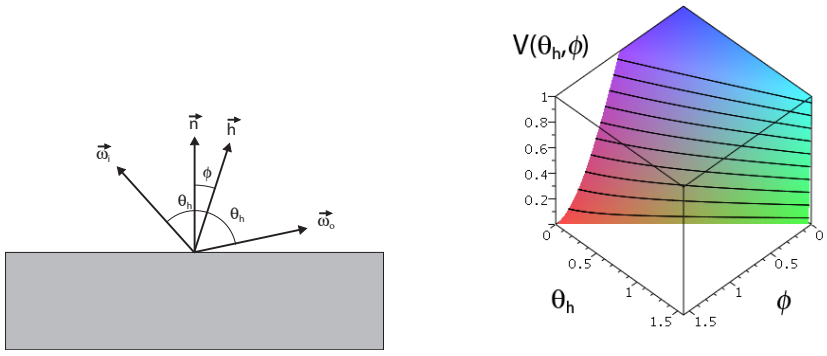
$$f_r(\vec{\omega}_i, \vec{\omega}_o) = k_s f_{r,s}(\vec{\omega}_i, \vec{\omega}_o) + (1 - k_s\rho(\vec{\omega}_o)) \cdot \frac{k_d}{\pi} \cdot (1 - k_s\rho(\vec{\omega}_i)) \tag{28}$$

As mentioned the term $(1 - k_s\rho(\vec{\omega}_i))$ is the fraction of light which did not reflect in a single bounce. Though the remainder includes light which has been reflected in multiple bounces we can think of the term as representing the amount which was transmitted into the material since both are modeled by a diffuse term. We can thus think of the scale by the factor k_d as representing an approximation for any subsequent loss due to internal absorption. The additional scale by $(1 - k_s\rho(\vec{\omega}_o))$, to obey the symmetry rule, is justified by the final step which transmits the light out of the material and into the direction $\vec{\omega}_o$ specifically. Note that since the internal scattering is assumed to be evenly distributed, given our diffuse component, then transmittance from $\vec{\omega}_o$ and into the material is equivalent to transmittance from the material and into the direction $\vec{\omega}_o$. Equation (28) is also given in papers such as [DJ05] and [dL07].

Though $f_{r,s}$ is position invariant the factors k_s and k_d can be position dependent. Furthermore, the factor ρ only depends on a single direction, which is two dimensional, and can be precomputed. If the material is assumed to have isotropic roughness then the input is only one dimensional $\vec{\omega}_i \bullet \vec{n}$.

3 Results

In this section the different terms in the BRDF, given by equation (18), are visualized and the influence these have on its characteristics is discussed. Specifically, it was noted by Torrance–Sparrow in [TS67] that off–specular peaks are observed, in measured data, for both metallic and nonmetallic surfaces. Some authors have attributed this to the shape of the Fresnel reflection curve which is supported by figure 4(d). However, as pointed out in [TS67], since most metals are highly reflective across a large range of wavelengths this does not explain the off–specular peaks found in measured data for metals. Though these peaks are generally weaker for metals these should still be accounted for in the model to achieve better accuracy. According to Torrance–Sparrow these peaks, for metals, should be attributed to shadowing and masking of light which is accounted for by the geometry term in his model given here by equation (25). Thus we expect to find this behavior confirmed by our test results.



(a) illustration

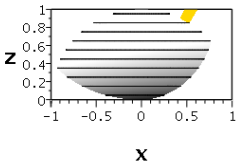
(b) visibility term

Figure 3: In the illustration shown in figure 3(a) all directions are in the same 2D plane. The angle θ_h is half the angle between $\vec{\omega}_o$ and $\vec{\omega}_i$ and the angle ϕ is the angle between $\vec{n}$ and $\vec{h}$. The geometry term is thus shown in figure 3(b) as a function of θ_h and ϕ.

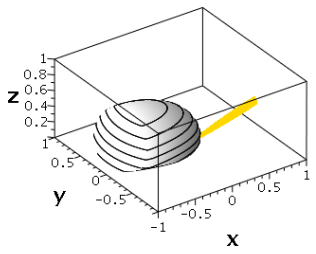
In equation (25) the minimum choice comes down to whether the incoming or the outgoing angle is larger relative to the normal $\vec{n}$. Let us assume for now that the outgoing angle is larger since a similar example is possible given the opposite scenario. We showed in section 2.5 that evaluation of this middle term in equation (25) can also be done after projection of $\vec{\omega}_o$ into the plane spanned by $\vec{n}$ and $\vec{h}$. Thus we can think of all three directions as given in the same plane as illustrated in figure 3(a). In this plane we get the following geometry term

$$V'(\vec{\omega}_o) = \min\left(1, \frac{2\cos(\phi)\cos(\phi+\theta_h)}{\cos(\theta_h)}\right)$$

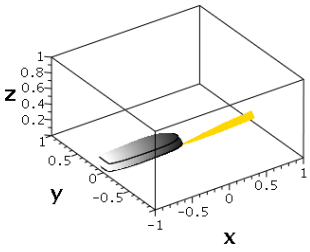
which is valid for $\phi + \theta_h \in \left]0; \frac{\pi}{2}\right[$ and is shown in figure 3(b). As pointed out in [TS67] if we combine this with the division by $\vec{n} \bullet \vec{\omega}_o$ in equation (18) then we do get a spike as θ_h approaches $\frac{\pi}{2}$ as we do for nonmetals.



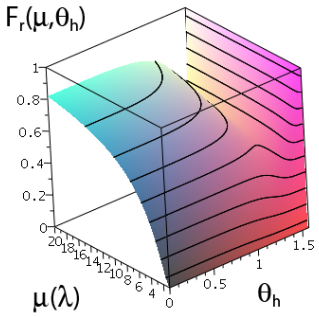
(a) ω_i at 27°



(b) ω_i at 75°



(c) ω_i at 86°



(d) Fresnel

Figure 4: In figures 4(a)–4(c) we see equation (25) shown, in each case, as a spherical function of $\vec{\omega}_o$ for a fixed incident direction $\vec{\omega}_i$. This direction is given in the XZ–plane by increasing angle to the normal $\vec{n}$ and is illustrated by a yellow spike. Figure 4(d) shows Fresnel reflection for unpolarized light as a function of a real valued index of refraction μ and the angle θ_h between $\vec{\omega}_o$ and $\vec{h}$. Note the peak as θ_h approaches $\frac{\pi}{2}$.

For a fixed incident direction the BRDF is a spherical function of the outgoing direction. To gain better insight into how the geometry term really affects the BRDF it is helpful to show it in such a frame. Figure 4(a) shows this at near normal incidence which causes $\vec{\omega}_o$ to dominate the minimum choice in equation

(25). Thus we see a fall-off occurring at shallow angles for the outgoing direction. Furthermore, the shape has a tilt which is due to the $\vec{n} \bullet \vec{h}$ term since the angle between the two is greater when $\vec{\omega}_o$ is on the same side of the normal $\vec{n}$ as the incident direction $\vec{\omega}_i$.

In figure 4(b) we see the visibility function as the direction of incidence is at a shallow angle. The reflection of the incident direction by the normal $\vec{\omega}'_o = R_{\vec{n}}(\vec{\omega}_i)$ represents the ideal outgoing direction where $\vec{n}$ and $\vec{h}$ are aligned. The visibility function (25) penalizes, due to the $\vec{n} \bullet \vec{h}$ term, when these are misaligned. Subsequently, in figure 4(b) we see a fall–off occurring as the outgoing direction begins to deviate from the ideal direction $\vec{\omega}'_o$. At an even more shallow incident angle, as is shown in figure 4(c), the fall–off becomes more rapid. Note that for such a grazing incident direction the half angle θ_h, relative to $\vec{\omega}'_o$, approaches $\frac{\pi}{2}$. The rapid fall–off that we see, in figure 4(c), when θ_h approaches $\frac{\pi}{2}$ also agrees with figure 3(b). Furthermore, the maximum value of Fresnel reflectance, when $\mu \in \mathbb{R}$, is also attained at $\theta_h = \frac{\pi}{2}$.

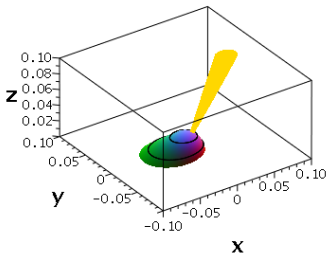
Next we will show the Fresnel reflectance function, again on the hemispherical domain, applied to two different materials. The first, methyl alcohol/methanol, is a wood derivative and has an index of refraction of approximately $\mu = 1.329$. The second, aluminum, is a metal which at a wavelength of $500nm$ has a complex index of refraction $\mu = (0.769, 6.08)$.

In figure 5(a) the Fresnel reflectance, at near normal incidence on methanol, resembles a moderate and rounded shape. In contrast we see in figure 5(b), at a shallow incident angle, the shape of a spike which can account for off–specular peaks. The same examples are shown for aluminum in figures 5(c) and 5(d) respectively. Unlike the methanol the Fresnel reflectance now appears to be almost constant for all outgoing directions. In figure 5(d) there is a very marginal peak near $\vec{\omega}'_o$ as θ_h approaches $\frac{\pi}{2}$. This agrees with the observations made by Torrance–Sparrow that Fresnel reflectance cannot account for the off–specular peaks observed in measured BRDF data for metals.

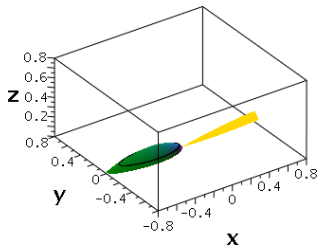
Finally, the last term in the BRDF equation (18) is the PDF $p(\vec{h})$. The choice of $p(\vec{h})$ is often made to be that of a normally distributed surface (see section 5.3 in [BS87]). Such a distribution yields an isotropic surface roughness since the distribution is radially symmetric about the normal. However, Torrance–Sparrow assumes this distribution to be given with respect to the angle between $\vec{n}$ and $\vec{h}$ as opposed to [BS87] and [CT81] who assume the distribution is based on displacements relative to the mean level $z = 0$. The argument for the latter is that this form of distribution provides a physical interpretation of the distribution parameter σ, which is the standard deviation of displacements, and makes fitting to the distribution easy. A broad base represents a rough surface as opposed to a thin distribution which corresponds to a near planar surface.

As mentioned in section 1 fitting is not the priority of this paper. Thus we proceed with a different PDF with a Gaussian–like behavior, known as the *von Mises–Fisher* distribution, but renormalized to the hemispherical domain.

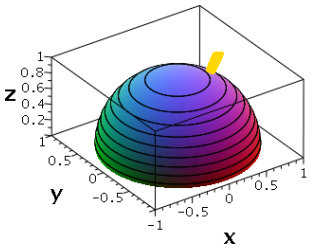
$$p(\vec{h}) = \frac{\kappa}{2\pi\left(e^{\kappa} - 1\right)} e^{\kappa(\vec{h} \bullet \vec{n})} \tag{29}$$



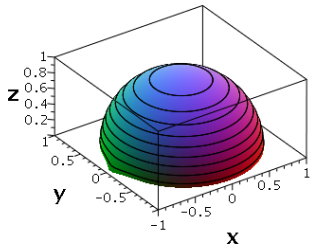
(a) methanol with ω_i at 27°



(b) methanol with ω_i at 86°



(c) aluminum with ω_i at 27°

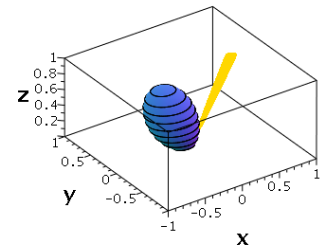


(d) aluminum with ω_i at 86°

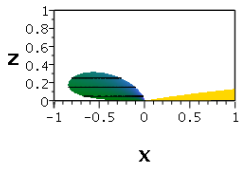
Figure 5: In figures 5(a)–5(d) we see, for all outgoing directions, the Fresnel reflectance as a function of θ_h. On the left side the light is at near normal incidence and on the right we have a shallow angle. The upper figures were generated using the index of refraction $\mu = 1.329$ for the liquid methanol. The lower figures were done for aluminum at a wavelength of 500 in which case $\mu = (0.769, 6.08)$ where the second component is the extinction coefficient.

We do this because it is easy to integrate, using importance sampling, with respect to this distribution since, as we shall see, it provides a straightforward inverse of the cumulative density function. A method to fit directional data to a mixture of such distributions is known as spherical EM [BDGS05]. A thorough derivation and description of this algorithm was also given in my master's thesis [Mik08] and the interested reader is referred to these papers.

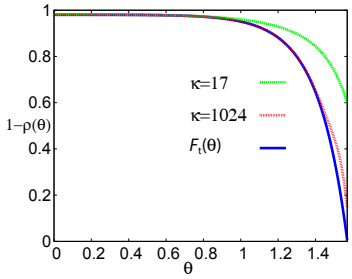
If we visualize $p(\vec{h})$ for all outgoing directions the result is a lobe around the,



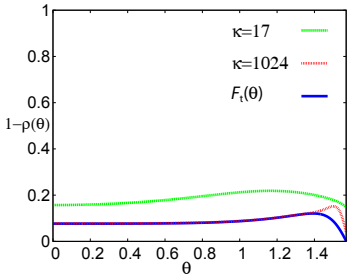
(a) $p(\vec{h})$ with ω_i at 27°



(b) $p(\vec{h})$ with ω_i at 86°



(c) methanol



(d) aluminum

Figure 6: Figures 6(a) and 6(b) show equation (29) at incident angles of 27° and 86° respectively. In figures 6(c) and 6(d) we see $1-\rho(\theta)$ and $F_t(\theta)$ for methanol and aluminum respectively. The parameter θ is the angle between $\vec{n}$ and $\vec{\omega}$. The function ρ is given by equation (27) and $F_t(\theta)$ is the Fresnel transmittance.

previously mentioned, ideal outgoing direction $\vec{\omega}_o'$. This is shown, for $\kappa = 17$, at near normal incidence in figure 6(a) and at low incidence in figure 6(b). As the value for κ is increased the lobe will become thinner and eventually approach the impulse function in the direction of $\vec{\omega}_o'$.

We can accelerate integration, with respect to $\vec{h}$, by importance sampling according to $p(\vec{h})$ in equation (18). However, outgoing radiance is determined using equation (3) which integrates with respect to $\vec{\omega}_i$. We solve this by substi-

tution using relation (15) which gives

$$\begin{aligned} L_o(\vec{\omega}_o) &= \int_{\Omega_{2\pi}} \delta\left(\vec{\omega}_i' \in \Omega_{2\pi}\right) f_r(\vec{\omega}_o, \vec{\omega}_i') L(\vec{\omega}_i')(\vec{\omega}_i' \bullet \vec{n}) \cdot 4(\vec{h} \bullet \vec{\omega}_o) d\vec{h} \\ &= \int_{\Omega_{2\pi}} \delta\left(\vec{\omega}_i' \in \Omega_{2\pi}\right) F_r(\vec{\omega}_o \bullet \vec{h}) L(\vec{\omega}_i') V' \frac{\vec{h} \bullet \vec{\omega}_o}{\vec{n} \bullet \vec{\omega}_o} \cdot \frac{p(\vec{h})}{\vec{n} \bullet \vec{h}} d\vec{h} \end{aligned} \tag{30}$$

where $\vec{\omega}_i' = R_{\vec{h}}\left(\vec{\omega}_o\right)$ and $V' = V'(\vec{h}, \vec{\omega}_o, \vec{\omega}_i')$. We have also used that the roles of $\vec{\omega}_i$ and $\vec{\omega}_o$ in relation (15) are interchangeable and that $\vec{h} \bullet \vec{\omega}_o = \vec{h} \bullet \vec{\omega}_i'$. Furthermore, we use $\delta\left(\vec{\omega}_i' \in \Omega_{2\pi}\right)$ to restrict the integral to incident directions above the tangent plane.

To integrate using importance sampling, with respect to some distribution p, an even distribution of probabilities is inverted into the corresponding set of sampling locations on the domain. An efficient way to do this is using the inverse of the cumulative density function (CDF). Thus we derive the inverse of the CDF for equation (29) in the following.

The PDF is defined on the surface of the hemisphere $\Omega_{2\pi}$. Let $\vec{h}$ be given in spherical coordinates $(\theta, \phi) \in \left[\frac{\pi}{2}; \pi\right] \times [0; 2\pi[$ such that $\frac{\pi}{2}$ at the horizontal plane and $\theta = \pi$ at the north pole. Since the distribution is isotropic $\phi = 0$ at some arbitrary axis in the tangent plane. The angle between $\vec{n}$ and $\vec{h}$ is $\pi - \theta$. Let $r \in \left[\frac{\pi}{2}; \pi\right]$ be a given value. Thus the CDF is equal to

$$\begin{aligned} P(r) &= \frac{\kappa}{2\pi\left(e^{\kappa} - 1\right)} \int_0^{2\pi} \int_{\frac{\pi}{2}}^{r} e^{\kappa \cos(\pi - \theta)} \sin\theta d\theta d\phi \\ &= \frac{\kappa}{e^{\kappa} - 1} \int_{\frac{\pi}{2}}^{r} e^{-\kappa \cos(\theta)} \sin\theta d\theta \end{aligned}$$

and represents the probability that $\vec{h}$ is chosen such that $\pi - \theta < r$. Next we apply substitution using $s = -\cos(\theta)$ and $ds = \sin\theta d\theta$ which gives

$$\begin{aligned} P(r) &= \frac{\kappa}{e^{\kappa} - 1} \int_0^{-\cos r} e^{\kappa s} ds \\ &= \frac{e^{-\kappa \cos r} - 1}{e^{\kappa} - 1} \end{aligned}$$

Let $t \in]0; 1]$ be the probability associated with r such that $t = P(r)$. The inverse of the CDF is thus equal to

$$\begin{aligned} r &= P^{-1}(t) \\ &= cos^{-1}\left(\frac{\ln\left(t \cdot e^{\kappa} + (1 - t)\right)}{-\kappa}\right) \end{aligned} \tag{31}$$

$$\tag{32}$$

The analytical CDF allows us to determine the solid angle $d\vec{h}$ associated with

each sample point using

$$\begin{aligned}\frac{dr}{dt} &= \frac{dP^{-1}(t)}{dt} \\ &= \frac{1}{\frac{dP(P^{-1}(t))}{dr}} \\ &= \frac{1}{\frac{dP(r)}{dr}} \\ &= \frac{e^{\kappa\cdot(\cos(r)+1)} - e^{\kappa\cdot\cos(r)}}{\kappa\sin(r)}\end{aligned}$$

Next for some sample (θ_0, ϕ_0) where $t_0 = P(\theta_0)$ and assuming there are N samples we get

$$\begin{aligned}d\vec{h} &= \sin\theta_0 d\theta d\phi \\ &= \sin\theta_0 \cdot \frac{2\pi\frac{dr(t_0)}{dt}}{N} \\ &= \frac{2\pi}{N} \cdot \frac{e^{\kappa\cdot(\cos(\theta_0)+1)} - e^{\kappa\cdot\cos(\theta_0)}}{\kappa}\end{aligned} \tag{33}$$

The parameters $\vec{\omega}_o$ and $\vec{\omega}_i$ are interchangeable in equation (15). Thus we can use it to determine the incident solid angle $d\vec{\omega}_i$ from, an outgoing direction $\vec{\omega}_o$, and the solid angle associated with $\vec{h}$ which can be determined using equation (33).

In section (2.6) we gave the BRDF equation (28) which includes the diffuse term to account for multiple bounces and subsurface scattering. In this equation we multiplied $1 - k_s\rho(\vec{\omega}_i)$ onto the diffuse term to preserve energy conservation and, subsequently, $1-k_s\rho(\vec{\omega}_o)$ also to preserve symmetry. It has been mentioned several times in this section that the specular peaks occur at shallow angles and so we expect the term $1 - k_s\rho(\vec{\omega})$ to decrease as the angle between $\vec{n}$ and $\vec{\omega}$ approaches $\frac{\pi}{2}$.

The equation (27), for $\rho(\vec{\omega})$, is equal to equation (3) with the radiance field $L(x, \omega') = 1$ for all ω'. Thus the variant, after substitution, given by equation (30) is applicable to $\rho(\vec{\omega})$. Since $p(\vec{h})$ approaches the impulse function $\delta\left(\vec{h} = \vec{n}\right)$ exponentially, for increasing κ, equation (30) gives us

$$\lim_{\kappa\to\infty} \rho(\vec{\omega}) = F_r(\vec{\omega} \bullet \vec{n})$$

Thus, for $k_s = 1$, we expect $1 - \rho(\vec{\omega})$ to approach the Fresnel transmittance $F_t(\vec{\omega} \bullet \vec{n})$. Since we have assumed an isotropic distribution ρ only depends on the angle θ between $\vec{n}$ and $\vec{\omega}$. This allows us to plot $1 - \rho(\theta)$, in 2D, for a chosen value of κ. Examples are shown in figures 6(c) and 6(d) for methanol and aluminum respectively. The green curve represents results for $\kappa = 17$, the red curve is for $\kappa = 1024$ and the blue curve is the Fresnel transmittance $F_t(\theta)$. As expected there is a decline as θ approaches $\frac{\pi}{2}$ and, as we see, increasing κ

does bring results closer to $F_t(\theta)$. An important difference between aluminum and methanol is that the significance of the diffuse term is very moderate for aluminum compared to methanol even at $\theta = 0$.

To see the effect, on visual appearance, we apply the BRDF model to a simple sphere using OpenGL. The goal is not to produce images which are, convincingly, similar to methanol and aluminum but, instead, to observe the impact on results. Thus the parameters $\kappa = 17$ and $\kappa = 1024$, that we use, are empirically chosen to represent a broad, and a thin, PDF respectively. We will, in the following, see how such choices affect the mix between the diffuse and specular term. To describe the field of incoming radiance $L(\vec{\omega}_i)$ we use an environment map which is freely available at [ICT]. The map only stores three channels across the spectral frequency domain which is common in computer graphics (red, green and blue). In our lighting model, it is possible, to produce more samples across a wide range of frequencies and convolve against the XYZ matching functions (see [oI04]). However, this level of accuracy is not necessary for our purpose so we will settle with the usual red, green and blue. Furthermore, we set the reflectance constants to $k_s = 1$ and $k_d = (0.08, 0.172, 1)$ where the latter, which is predominantly blue, is chosen to amplify the distinction, in results, between the specular and diffuse term.

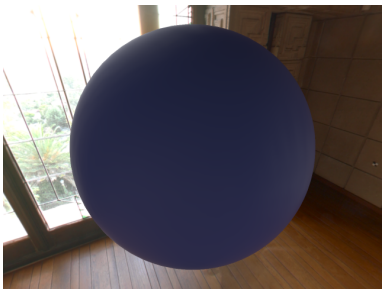
The result for aluminum with $\kappa = 1024$ is shown in figure 7(b). Given figure 6(d) and equation (28) the role of the diffuse term should be marginal which is confirmed since there is no sign of color modulation due to k_d. The reflection is blurred due to the moderate distribution thickness $\kappa = 1024$. In figure 7(d) the same test is shown but with a broad distribution $\kappa = 17$ which, then, heavily blurs the reflection.

The transmittance profile 6(c) for methanol is above 80% until θ is greater than approximately $\frac{5\pi}{12}$ (ie., 75° in degrees). As a consequence, in figure 7(a), where $\kappa = 1024$, the diffuse term dominates the interior of the sphere. Conversely, as θ approaches $\frac{\pi}{2}$ the profile approaches 14%. Subsequently, at the silhouette, as seen from the observer, the diffuse term is significantly reduced, which allows the specular term to dominate, such that color modulation is minimal. This is due to the scale by $1-\rho(\vec{\omega}_o)$, in equation (28), and the effect agrees with the observation made by Torrance–Sparrow that diffusely reflected light is minimal at shallow angles. In comparison the traditional Phong lighting model would mispredict this given that the accumulated diffuse term is independent of the view direction $\vec{\omega}_o$.

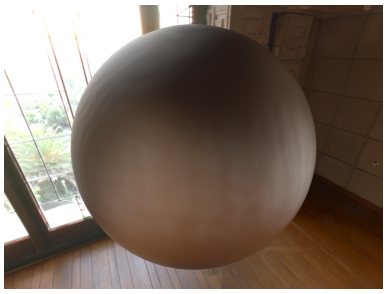
In figure 7(c) we see the same test for methanol but using the broad distribution $\kappa = 17$. As we see the specular reflection at the silhouette is now very subtle. This is because the transmittance profile in figure 6(c) associated with $\kappa = 17$ stays above 60%. Furthermore, since the distribution of microfacet normals is broad this makes the reflected direction less uniform which reduces the outgoing density.

(a) methanol, $\kappa = 1024$

(b) aluminum, $\kappa = 1024$



(c) methanol, $\kappa = 17$



(d) aluminum, $\kappa = 17$

Figure 7: The material properties of methanol were used to render the images to the left and those of aluminum were used to produce those on the right side. Furthermore, a broad distribution of microfacet normals was used in figures 7(c)-7(d) which creates a heavy blur. In figures 7(a)-7(b) a, moderately, thin distribution was used.

4 Conclusion

In this paper we have given a derivation of the analytical, physically based BRDF, by Torrance–Sparrow [TS67]. We have done this starting from an explicit formulation of outgoing radiance resulting from incident flux reflected at a small surface comprised of microfacets at an infinitesimal scale.

A clarification at this level of detail was not given in the original paper [TS67]. Furthermore, the function known, today in computer graphics, as the Torrance–Sparrow BRDF is not given in the same form in the original paper [TS67]. Thus our derivation here provides a justification for the form used today.

During our analysis we made the observation that the mutual visibility term must be approximated by a local evaluation to complete a substitution step which, ultimately, leads to the analytical BRDF. This subtlety is interesting since visibility is known to be a global problem. The same term in [TS67] is referred to as a geometry term and is discussed in a different context, which is, modeling off–specular peaks observed in measured data. In our work we justify this term mathematically instead.

The geometry term in [TS67] is expressed as a function of parameters such as cavity size and occlusion coverage. In computer graphics it is more practical to express the function in terms of dot products between vectors which are typically known in this context such as: incident direction, observer direction and surface normal. Thus Jim Blinn derived such a formulation of the geometry term in his paper [Bli77] using many steps of trigonometric identities. In our paper we have given a different derivation of the term which is simpler and more intuitive. The geometry term is based on the assumption that the surface is a plane full of cavities. For future work it might be interesting to explore other configurations such as the saw-tooth surface. This would be a good match for hair fibers given that these are composed of tilted cuticle scales.

It is discussed in [TS67] that subsurface scattering and light reflected in multiple bounces is accounted for by a simple diffuse term. However, the issue of energy conservation and symmetry is never discussed. Furthermore, the Torrance–Sparrow BRDF used today in graphics only accounts for light reflected in a single bounce, at the surface, and as we have mentioned does not exist in the same form in [TS67]. This prompts the question of how to integrate the diffuse term with today's model. We suggest and explain a model which has been used by some other authors but to our knowledge has never been explained in published work.

Acknowledgments.

This author would like to thank Swaminathan Narayanan and Kasper H. Nielsen for helpful discussions and Kenny Erleben for additional constructive comments and proof reading.

References

[AP07] Michael Ashikhmin and Simon Premoze. Distribution–based brdfs. Technical report, University of Utah, March 2007. Available at http://www.cs.utah.edu/~premoze/dbrdf/.

[BDGS05] Arindam Banerjee, Inderjit S. Dhillon, Joydeep Ghosh, and Suvrit Sra. Clustering on the unit hypersphere using von mises-fisher distributions. *J. Mach. Learn. Res.*, 6:1345–1382, 2005.

[Bli77] James F. Blinn. Models of light reflection for computer synthesized pictures. *SIGGRAPH Comput. Graph.*, 11(2):192–198, 1977.

[BS87] Petr Beckmann and Andre Spizzichino. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Artech Print on Demand., 1987.

[CT81] Robert L. Cook and Kenneth E. Torrance. A reflectance model for computer graphics. In *SIGGRAPH '81: Proceedings of the 8th annual conference on Computer graphics and interactive techniques*, pages 307–316, New York, NY, USA, 1981. ACM.

[DJ05] Craig Donner and Henrik Wann Jensen. Light diffusion in multilayered translucent materials. *ACM Trans. Graph.*, 24(3):1032–1039, 2005.

[dL07] Eugene d'Eon and David Luebke. *GPU Gems 3*, chapter 14, pages 293–347. Addison-Wesley Publishing, 2007.

[Fl28] Siegfried Flgge. *Handbuch der Physik*, volume 20. Springer-Verlag., Berlin, Germany, 1928.

[GTS+97] Donald P. Greenberg, Kenneth E. Torrance, Peter Shirley, James Arvo, Eric Lafortune, James A. Ferwerda, Bruce Walter, Ben Trumbore, Sumanta Pattanaik, and Sing-Choong Foo. A framework for realistic image synthesis. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 477–494, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[ICT] High–resolution light probe image gallery. Available at http://gl.ict.usc.edu/Data/HighResProbes/.

[Ish78] Akira Ishimaru. *Wave Propagation and Scattering in Random Media*. Oxford University Press, 1978.

[JMLH01] Henrik Wann Jensen, Stephen R. Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. In *SIGGRAPH '01: Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 511–518, New York, NY, USA, 2001. ACM.

[Mik08] Morten S. Mikkelsen. Simulation of wrinkled surfaces revisited. Master's thesis, Department of Computer Science at the University of Copenhagen, 2008.

[MPBM03] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. In *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers*, pages 759–769, New York, NY, USA, 2003. ACM.

[MWLT00] Stephen R. Marschner, Stephen H. Westin, Eric P. F. Lafortune, and Kenneth E. Torrance. Image-based bidirectional reflectance distribution function measurement. *Appl. Opt.*, 39:2592–2600, 2000.

[NIK92] Shree K. Nayar, Katsushi Ikeuchi, and Takeo Kanade. *Surface reflection: physical and geometrical perspectives*, pages 60–83. Jones and Bartlett Publishers, Inc., , USA, 1992.

[NRH+77] F. Nicodemus, J. Richmond, J. Hsia, I. Ginsberg, and T. Limperis. Geometric considerations and nomenclature for reflectance. *National Bureau of Standards (US)*, 1977.

[oI04] CIE International Commission on Illumination. Colorimetry. Technical report, CIE Central Bureau, 2004.

[PH04] Matt Pharr and Greg Humphreys. *Physically Based Rendering: From Theory to Implementation.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.

[Pho75] Bui Tuong Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975.

[Sch94] Christophe Schlick. An inexpensive brdf model for physically-based rendering. *Eurographics, Computer Graphics Forum*, 13(3):233–246, 1994.

[TS67] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.*, 57(9), 1967.

[War92] Gregory J. Ward. Measuring and modeling anisotropic reflection. *SIGGRAPH Comput. Graph.*, 26(2):265–272, 1992.