Abstract

The bidirectional reflectance distribution function (BRDF) is a measure of the amount of light scattered by some medium from one direction into another. Integrating it over specified incident and reflected solid angles defines the reflectance, which can be easily related to the absorptance (or emissivity) of a sample. The BRDF can thus be taken as a fundamental quantity for the optical characterization of an object and it correspondingly is important in a large variety of applications. When a beam of electromagnetic radiation (visible, infrared, radar, etc.) strikes a body, it can scatter off the top or first surface, as well as from the volume or subsequent interfaces. However, the total amount of light reflected from the first surface depends primarily on the complex index of refraction of the illuminated medium (relative to that of the incident medium) and is often sufficiently large that this dominates the scattering from the material. On the other hand, the topography of this interface determines the angular distribution of the scattered radiation—smooth surfaces reflect almost entirely into the specular direction, while with increasing roughness the light tends to diffract into all possible directions. Ultimately an object will appear equally bright throughout the outgoing hemisphere if its surface is perfectly diffuse (i.e., Lambertian). Measuring and modeling the BRDF can thus give valuable information about the nature of a target sample.

This paper is organized as follows. The introduction lists some domains of study in which reflectance plays an important role, as well as the relationship between it and other quantities of importance in optics. Next, the nomenclature needed to define and characterize the BRDF is presented, along with some related issues. Following this, a long section reviews the principal analytical and numerical models used to describe first-surface scattering; this comprises the heart of the present paper and considerable effort has been expended to unify the often disparate notations and points of view in the literature. Finally, this report ends with a summary of select experimental measurements, most of which are quite recent and varied in style and purpose; this gives a flavor for the ongoing efforts in this field.

Introduction

A variety of models and measurements of reflectance have been described in the literature but their range of validity is generally restricted by the domain of interest of the authors. Roughly speaking, these domains can be described as follows:

- Remote sensing—aircraft or satellite measurements of terrestial vegetation and geography, usually under ambient atmospheric lighting conditions;
- Astronomy—telescopic measurements of planets and moons in the solar system illuminated directly or indirectly by the sun;
- Computer graphics—development of techniques intended to photorealistically simulate the surface appearance of a wide variety of materials;
- Military/commercial—field measurements under carefully controlled conditions of man-made objects, typically painted or bare-metal air or ground based targets;
- Optics—goniometric laboratory measurements of small, well-characterized samples.

Select models and experiments applicable to each of these five domains are discussed in this paper.

In general, theoretical models begin by considering light to be incident on a spot on the surface of an object from a range of solid angles $\Omega$ (which can vary from a delta function for a fully directional beam to $2\pi$ for the hemisphere) centered on polar (zenith) angle $\theta_i$ (defined with respect to the macroscopic surface normal) and azimuthal angle $\phi_i$ (conveniently defined with respect to some sample feature in the case of an anisotropic surface or arbitrarily and irrelevantly
for an isotropic one). This light must be either transmitted, absorbed, or reflected by the object; the fraction of the incident flux $P$ (radiant power in W) which is subsumed by each of these mechanisms is specified by the dimensionless ratios called the transmittance $\tau$, the absorptivity (or absorptance) $\alpha$ (not to be confused with the absorption coefficient given by $-\ln[1-\alpha]/L$ where $L$ is the sample length), and the reflectance $\rho$, respectively. I will always assume that the illuminated object is opaque, so that $\tau = 0$; accordingly, $\alpha + \rho = 1$. The reflected light (which can be more generally described as scattered light) is collected by a detector spanning a solid angle $\Omega$, centered on angles $\theta$ and $\phi$. The emissivity $\epsilon$ of the sample is defined to be the dimensionless ratio of the total radiant flux emitted by the sample to that of a blackbody having the same geometry and temperature. Kirchhoff’s law says that $\epsilon = \alpha$, as follows by imagining the sample to be enclosed inside an opaque cavity whose temperature is equal to that of the sample: this equality means that there can be no net gain from or loss to the blackbody environment by the object. Furthermore, by interposing narrowband filters between the sample and cavity, it is evident that the spectral emissivity $\epsilon(\lambda)$, defined in terms of the ratio of emitted fluxes in a unit wavelength interval centered around $\lambda$, must equal the spectral absorptivity $\alpha(\lambda)$, whose definition is similarly a ratio of spectral fluxes. [Note that I have eschewed the notation $\epsilon_s$, to avoid confusion with the spectral emission coefficient defined in Sec. 13.1.1 of Hecht (1998).] An object is said to be a graybody if $\epsilon(\lambda)$ is a constant less than 1. Nicodemus (1965, 1970) has argued that equality also holds for the directional quantities, $\epsilon(\theta, \phi) = \alpha(\theta, \phi)$, for emission into some direction and absorption from the same direction. [One could then call $\epsilon$ the hemispherical emissivity, equal to the average of the directional emissivity over all projected solid angles, analogous to Eq. (2) below.] Thus the directional and spectral dependence of the emitted light can be related to the reflectance and temperature (via the Stefan-Boltzmann law) of the sample. Furthermore the polarization of the emission is seen to be the complement of that of the reflected radiation at the same wavelength, a result which can also be understood by visualizing an emitting center as lying slightly below the surface of the sample, so that the light escapes in accordance with the Fresnel relations [cf. Eq. (6) below]; in fact, Sandus (1965) has argued that this holds even for emitting centers lying on the surface. Consequently, I will mainly focus attention on the reflectance for the remainder of this paper, even in the infrared region where the sample emission is significant at the ambient temperatures of interest.

**Reflectance Nomenclature**

The incident flux $P$ per unit illuminated area $A$ of an object is known as the irradiance (or illuminance) $H \equiv dP/dA$ in units of W/m$^2$; the corresponding emitted or scattered quantity is called the emittance or exitance. [Note that the element $dA$ of area is taken to be small on the macroscopic scales of interest, but large compared to the sub-resolution irregularities in the sample surface, and is taken to be perpendicular to the macroscopic surface normal (Nicodemus, 1965).] For simplicity, both the illuminance and the exitance will be referred to as the irradiance of the surface, to be distinguished from the intensity of the irradiating beam, $I \equiv dP/(dA \cos \theta)$, where the $\cos \theta$ term projects the element of surface area $dA$ into the direction of propagation of the beam, which is inclined at polar angle $\theta$ to the (macroscopic) surface normal. The irradiance per unit projected solid angle is called the radiance (or luminance, which corresponds to the photometric concept of brightness), $L \equiv dH/(\cos \theta \, d\Omega) = dI/d\Omega$ in units of W m$^{-2}$ sr$^{-1}$. Here the element of solid angle is $d\Omega \equiv \sin \theta \, d\theta \, d\phi$, sometimes (Nicodemus et al., 1977) a separate symbol is also introduced for the element of projected solid angle $\cos \theta \, d\Omega$, though I will not do so, to keep things simpler. Let the subscript $i$ refer to incident quantities and $r$ to reflected (scattered) terms. The bidirectional reflectance distribution function (BRDF) can now be defined as $f(\theta_i, \phi_i; \theta_r, \phi_r) \equiv dL_i/dH_r = dP_r/(dP_i \cos \theta, d\Omega_r)$ with units of sr$^{-1}$; here $dP_i$ is the incident power illuminating $dA$ from direction $(\theta_i, \phi_i)$ and $dP_r$ is the radiant power reflected from $dA$ into the outgoing solid angle $d\Omega_r$, centered about the direction $(\theta_r, \phi_r)$. Usually the surface is taken to be (macroscopically) planar and the incident beam to have a spatially uniform intensity profile, so
that one can integrate over a macroscopic area \( A \) of the sample to get \( f = \frac{dP_i}{(P_i \cos \theta_i) d\Omega_r} \) where now \( P_i \) is the incident power illuminating \( A \) and \( dP_i \) is the flux reflected from \( A \) into \( d\Omega_r \). For an isotropic surface, \( f \) is a function only of \( \phi_i - \phi_r \) and not of the two azimuthal angles individually. Unlike reflectances, which cannot exceed unity, the BRDF can be very large, becoming unbounded for a purely specular reflector—cf. Eq. (5) below. When necessary, a double subscript can be added to \( f \) to denote the polarizations of first the source and second the detector relative to the planes of incidence and reflection, respectively. In that case, these four values of \( f \) can be thought of as the elements of a \( 2 \times 2 \) matrix \( f \) and the component intensities of the incident and reflected radiation as two-vectors such that \( dL_r = f dL_i \). However, this description does not suffice to determine the scattering for other states of incident and reflected polarization, such as circular, because the phase information of the fields has not been retained. One approach to circumventing this limitation is to work instead with the (complex) components of the incident and reflected electric fields, \( E_i \) and \( E_r \), respectively, which are connected by a \( 2 \times 2 \) scattering matrix \( S \) (i.e., \( E_r = S E_i \)), which is known as the Jones calculus (Ruck et al., 1970). Alternatively, the Mueller matrix and Stokes vector representation can be used, wherein \( f \) is expressed as a \( 4 \times 4 \) matrix and the beams as intensity 4-vectors (Flynn and Alexander, 1995). This latter approach is needed for describing partially polarized radiation, while the Jones notation is used for coherent illumination and reflection. If the radiation is partially coherent, then a treatment in terms of the coherency (or polarization) matrix is necessary. For a nice discussion of the relation between the Jones, Mueller, and coherency matrices, with particular reference to the ensemble averaging necessary for handling scattering from statistical media, see Kim et al. (1987).

By integrating the BRDF over all scattered angles, one gets the dimensionless directional-hemispherical reflectance

\[
\rho(\theta, \phi; 2\pi) = \int_{2\pi} f(\theta_i, \phi_i; \theta_r, \phi_r) \cos \theta_i d\Omega_i,
\]

which gives the fractional amount of flux reflected into the entire hemisphere out of that incident from a particular direction. (The integral over \( 2\pi \) is of course a shorthand meaning integrate \( \theta \) from 0 to \( \pi/2 \) and \( \phi \) from 0 to \( 2\pi \), for a total solid angle of \( 2\pi \).) If instead one averages over all incident angles, the hemispherical-directional reflectance is obtained as

\[
d\rho(2\pi; \theta_r, \phi_r) = \frac{\cos \theta_i d\Omega_i}{\pi} \int_{2\pi} f(\theta_i, \phi_i; \theta_r, \phi_r) \cos \theta_i d\Omega_i.
\]

Note that the denominator is \( \pi \) and not \( 2\pi \) because of the \( \cos \theta \) projection factor. It should be clear from inspection of these two equations how to write other quantities of interest, such as the conical-hemispherical reflectance, the bihemispherical reflectance (or albedo), and so on—Nicodemus et al. (1977) tabulate a variety of such entities. If the prefactor before the integral in Eq. (2) is left off, the resulting quantity is instead called the hemispherical-directional reflectance factor, which is defined as the ratio of the flux reflected by the sample to that which would be reflected by an ideal Lambertian reflector for the same geometry, because \( f = 1/\pi \) for the latter according to the discussion following Eq. (9).

From the definition of the BRDF, the scattered radiance can be related to the incident radiance by

\[
L_r(\theta, \phi) = \int_{\Omega_i} L_i(\theta_i, \phi_i) f(\theta_i, \phi_i; \theta_r, \phi_r) \cos \theta_i d\Omega_i,
\]

which can be rather complicated to evaluate analytically. The emissivity is given by Kirchhoff’s law as

\[
e(\theta_i, \phi_i) = 1 - \int_{2\pi} f(\theta_i, \phi_i; \theta_r, \phi_r) \cos \theta_i d\Omega_r
\]

and is thus simply related to the (directional-hemispherical) reflectance, as discussed in the introduction.
According to the Helmholtz reciprocity theorem, \( f(\theta_i, \phi_i; \theta_r, \phi_r) = f(\theta_r, \phi_r; \theta_i, \phi_i) \), so that either direction may be that of the incident beam with the other the reflected beam. Note from Eqs. (1) and (2) that this would imply that the directional-hemispherical reflectance equals the hemispherical-directional reflectance factor; interestingly enough, however, equality of these two reflectance terms remains true even when Helmholtz reciprocity does not (Nicodemus, 1970). There has been considerable discussion about the conditions under which reciprocity holds (Clarke and Parry, 1985; Kriebel, 1996; Shirley et al., 1997; Snyder, 1998)—for example, it clearly does not apply to a Faraday isolator. It is generally accepted however that the theorem is obeyed by most materials (including compound objects) under ordinary conditions and hence models which are not reciprocal are generally considered unphysical and experimental failures are often attributed to measurement errors or limitations.

BRDF Models of Surface Scattering

It is often possible to distinguish surface scattering from volume scattering. Some samples (e.g., bare metals) have a penetration depth (equal to the reciprocal of the absorption coefficient) which is so small that all of the reflection can be assumed to occur at its illuminated or first surface. Other samples (e.g., a painted object) exhibit scattering both from the first surface (and possibly subsequent surfaces, such as the paint-substrate interface) as well as from the bulk interior (e.g., due to defects or paint pigments). Still other samples (e.g., a forest imaged from an airplane) do not have a well-defined first surface at all and could perhaps be best described as pure volume scatterers (Snyder and Wan, 1998). Volumetric reflection almost invariably involves multiple scattering (although a sufficiently cratered surface can as well). This greatly adds to the complexity of the problem and is not fully understood. For this reason, I will restrict attention to first-surface scattering.

Begin by considering the simplest case. A purely specular reflector (i.e., a perfectly smooth and planar mirror) has a BRDF equal to zero unless \( \theta_r = \theta_i \) and \( \phi_r = \phi_i + \pi \), in which case the reflectance is described by the Fresnel equations, so that (Ellis, 1994)

\[
f = R(\theta_i) \frac{\delta(\theta_r - \theta_i)\delta(\phi_r - \phi_i - \pi)}{\cos \theta_r \sin \theta_i},
\]

assuming that the illuminated area of the mirror is large enough (compared to the wavelength of the light) that diffraction from it into off-specular directions is negligible. For nonmagnetic materials (Hecht, 1998),

\[
R_s = |r_s|^2 \quad \text{with} \quad r_s = \frac{\cos \theta_i - \sqrt{n^2 - \sin^2 \theta_i}}{\cos \theta_i + \sqrt{n^2 - \sin^2 \theta_i}}
\]

and

\[
R_p = |r_p|^2 \quad \text{with} \quad r_p = \frac{n^2 \cos \theta_i - \sqrt{n^2 - \sin^2 \theta_i}}{n^2 \cos \theta_i + \sqrt{n^2 - \sin^2 \theta_i}}
\]

for polarization perpendicular (“senkrecht,” also called horizontal or TE and denoted by an \( h, - \), or \( \perp \)) and parallel (sometimes referred to as vertical or TM and symbolized by \( v, + \), or \( \parallel \)) to the plane of incidence, respectively. Here \( n = n_i/n_t \) where \( n_i \) and \( n_t \) are the (complex) refractive indices of the incident medium and the transmitted medium (i.e., the top layer of the sample), respectively; \( n_i = 1 \) and \( n_t = n \) correspond to the usual case of light in space scattering off a medium of index \( n \). Any incident beam can of course be decomposed into \( s \) and \( p \) components, provided that proper accounting is made of their relative phase. Recall that these Fresnel reflectances are strongly dependent upon the angle of incidence—for example, \( R_s \) falls to zero at the Brewster angle for a dielectric (or has a minimum at the pseudo-Brewster angle for a conducting medium) but rises to unity for grazing incidence. Substitution of Eq. (5) into Eq. (1)
identifies $R(\theta)$ as the directional-hemispherical reflectance; if the mirror is ideal (i.e., has infinite conductivity) then this reflectance is unity for any polarization and wavelength (Born and Wolf, 1965). Note that a flat specular reflector does not depolarize incident $s$ or $p$ radiation, although it will reverse the handedness of circular light and rotate the plane of polarization of light linearly polarized in directions other than $s$ and $p$, as discussed in more detail later.

All real surfaces have some roughness however. If the characteristic size scales (to be defined more precisely below) of this roughness are large compared to the wavelength of the light, then the approximation of geometrical optics holds. In that case, the simplest model for the surface scattering is that of Torrance and Sparrow (1967), who pictured the surface as being comprised of small, randomly oriented, specular facets. Neglecting shadowing and masking of the facets by each other, the BRDF for single scattering from this isotropic surface is easily shown to be

$$f = R(\beta) \frac{\Xi(\theta_n)}{4\cos\theta_i \cos\theta_r}, \quad (7)$$

where $2\beta$ is the angle between the incident and reflected directions, known as the bistatic angle ($2\beta = 0$ corresponds to monostatic or back scattering, while $2\beta = \pi$ refers to forward scattering pertinent to the optical extinction theorem for example), and $\Xi(\theta_n)$ is the density function (in units of sr$^{-1}$) of facet normals pointing in the direction $\theta_n$ relative to the macroscopic surface normal, i.e., $\Xi(\theta_n) d\Omega_n dA$ is the total (not the projected) surface area of the facets spanned by $dA$ which have normals lying within solid angle $d\Omega_n$ (Maxwell and Weiner, 1974). Clearly this density function must be normalized such that

$$\frac{\pi}{2} \int_0^{\pi/2} \Xi(\theta_n) \cos\theta_n \sin\theta_n d\theta_n = 1. \quad (8)$$

Torrance and Sparrow took $\Xi$ to be a Gaussian distribution function with zero mean. It can be shown that $\cos 2\beta = \cos\theta_i \cos\theta_r + \sin\theta_i \sin\theta_r \cos(\phi_i - \phi_r)$, $\cos\theta_n = (\cos\theta_i + \cos\theta_r)/2\cos\beta$, and $d\Omega_n = d\Omega/4\cos\beta$ using spherical geometry. Observe that Eq. (7) explicitly satisfies Helmholtz reciprocity. It can be rewritten in terms of experimentally measurable quantities by noting that $\Xi(\theta_n)$ is proportional to the monostatic BRDF $f(\theta_i,0;\theta_r,0)$. The polarization dependence is determined solely by the Fresnel coefficients; hence, in the plane of incidence there are only 4 independent Mueller matrix elements—e.g., the 11, 12, 33, and 34 terms, which can be related to the real and imaginary parts of the 2 diagonal elements, $r_s(\beta)$ and $r_p(\beta)$, of the Jones scattering matrix. Videen et al. (1992) have shown that this simple polarization prediction is in quite good agreement with experimental measurements on scratched or sandblasted copper and aluminum. Sung and Eberhardt (1978) have improved on the model by allowing the facets to be slightly curved and calculating a perturbative correction using the Rayleigh-Rice method described later. The above BRDF needs to be multiplied by a geometric attenuation factor $G(\hat{i},\hat{n},\hat{r})$ to correct for shadowing and masking, the former referring to the partial illumination of a facet shadowed by an adjacent one and the latter to the partial visibility of a facet occluded by another. Here $\hat{i}$, $\hat{n}$, and $\hat{r}$ refer to unit vectors in the source, facet normal, and detector directions respectively, as seen from the illuminated spot; note that $2\beta = \cos^2(\hat{i} \cdot \hat{r})$. For simplicity, Torrance and Sparrow supposed that each facet forms the side of a symmetric V-groove cavity whose axes are parallel to the macroscopic surface (but have random azimuthal orientations) and whose upper edges all lie in the same plane. In that case, the fraction $G$ of a given facet surface which contributes to the reflected flux (i.e., which is both illuminated and visible) can be found geometrically in terms of a range of unobscured angles of the projections of $\hat{i}$ and $\hat{r}$ into the plane defined by $\hat{n}$ and the macroscopic surface normal; the grooves are assumed to be long enough that end effects can be neglected. Although this model for the geometric attenuation is very simple, it successfully predicts that for large angles of incidence the peak in the BRDF in the plane of incidence occurs at an angle of reflection larger than specular. Such off-specular peaks are seen experimentally, both for coarsely roughened metals and dielectrics.
Finally, Torrance and Sparrow assumed that the multiply reflected light (and any volume scattering) was purely diffused. A purely diffuse (Lambertian) scatterer has two properties. First, its reflected radiance \( L_r \) is independent of the viewing angle and hence is equal to the hemispherical exitance \( H_r \), divided by \( \pi \) (again not \( 2\pi \)). The reflected flux per unit solid angle is proportional to \( \cos \theta_r \), which is called Lambert’s law; since the projected area of a surface element also varies as \( \cos \theta_r \), this is consistent with the fact that \( L_r \) is independent of the viewing angle. Equation (3) then implies that \( f \) is independent of \( \theta_r \) and \( \phi_r \), and so Eq. (1) becomes

\[
f = \frac{\rho(\theta_r,\phi_r;2\pi)}{\pi}.
\]

Second, its directional-hemispherical reflectance is independent of the angle of incidence and thus is equal to the bihemispherical reflectance; this property is a statement of reciprocity. The BRDF is then related to the bihemispherical reflectance according to

\[
f = \frac{\rho(2\pi;2\pi)}{\pi}
\]

and is independent of both the incident and reflected angles. If the surface is ideal (i.e., has unit albedo) then the BRDF further simplifies to \( 1/\pi \). Incidentally, note that a blackbody has zero albedo according to Eq. (4) and hence obeys Lambert’s law in emission; this explains why the sun appears as a uniformly bright disk even though it actually is a sphere. It should be emphasized that Eq. (9) does not imply that \( L_r \) is independent of the incident angle: on the contrary, Eq. (3) says that the reflected radiance due to a well-collimated incident beam of fixed radiance is proportional to \( \cos \theta_i \) because the flux incident on a unit area of the surface varies in this way (Torrance and Sparrow, 1967); this means that a spherical Lambertian scatterer illuminated from a fixed direction will appear shaded around the edges (Oren and Nayar, 1995), not flat as some authors have incorrectly claimed (Ellis, 1994), in striking contrast to a spherical Lambertian emitter. It should also be borne in mind that many rough objects, such as an ordinary piece of paper, are reasonably Lambertian at near-normal viewing angles, but become increasingly specular at grazing angles due to the increase in the first-surface Fresnel reflectance as well as to the compression of the surface roughness in the viewing direction—this latter effect gives rise, for example, to the horizontally streaked reflections commonly seen off waxed tile floors (Shirley et al., 1997). Finally, note that by substituting Eq. (5) or (9) into (1) or (2), the directional-hemispherical reflectances of an ideal mirror and of an ideal diffuser are found to be equal, and likewise for the hemispherical-directional reflectances (Judd, 1967).

Multiple scattering from the surface or volume of an object tends to depolarize the incident light; even two reflections from a valley can cross \( s \) or \( p \) polarization, as nicely illustrated in Fig. 14 of O’Donnell and Mendez (1987). For this reason, the light scattered by a Lambertian surface is typically assumed to have random polarization. However, this is not necessarily true, because the two defining properties which led to Eq. (9) do not explicitly require multiple scattering. For example, one could imagine choosing the density function in Eq. (7) to give a BRDF that is at least approximately independent of incident and reflected angles, thus characterizing a Lambertian diffuser, although only singly scattered, specular reflections from isotropic facets are involved which consequently cannot depolarize \( s \) or \( p \) light. Even volume scattering from a paint does not fully randomize the polarization of the radiation, as Ellis (1996b) has demonstrated experimentally. Therefore, one should be careful about assuming that diffusers always reduce the degree of polarization of a beam to nearly zero, in the absence of detailed knowledge of the scattering processes.

Oren and Nayar (1995) have developed an alternative form of Torrance and Sparrow’s model in which each facet is assumed to reflect purely diffusely rather than specularly. In effect, this presumes that the surface has two distinct roughness scales and hence is said to be compositely rough (Barrick, 1970): a coarse range specified by macroscopically flat facets whose dimensions are large compared to the wavelength of the light, and a fine range responsible for Lambertian scattering from each facet. [A very different two-roughness-scale model has been developed by Leader (1979).] Their model was developed to describe the reflectances of materials such as plaster, sandpaper, and cloth for computer rendering applications, and they performed
experimental measurements which gave fairly good agreement with the calculations. Although one might suppose that the overall effect of the two roughness scales in the model is merely to further roughen the surface and hence leave the overall scattering approximately Lambertian, this is incorrect: the reflectance is markedly non-Lambertian and in fact has a strong backscattering peak for large angles of incidence which gets cut off for reflection angles beyond the source direction (i.e., \( \theta_s < -\theta_i \)). Notice that this is completely opposite to the strong forward scattering peak predicted by the Torrance and Sparrow model.

The cause of the backscattering peak is reminiscent of Hapke’s (1963) shadow-hiding explanation for the opposition effect of the moon, wherein its brightness peaks at full moon when the sun is directly behind the earth. In the lunar case, the relevant scattering is volumetric, because the dust on the moon is very porous and is taken to be an open network of particles. Particles closer to the surface cast shadows down on the lower particles, which thus reduces the reflected radiance in every direction except the incident one where the shadows are hidden by the particles that created them. To put it another way, the light will be attenuated both as it penetrates the medium and as it leaves it after being reflected, unless it backscatters into the incident direction in which case it can retrace its path out and escape without being blocked. It is assumed that the albedo \( \rho \) of the individual particles is small enough that multiple scattering is negligible and that the particles are large enough that diffraction around their edges is minimal, since either of these would tend to wash out the effect. After making a number of simplifying approximations, the BRDF of the medium can then be expressed as the product of three factors, each of which satisfies Helmholtz reciprocity,

\[
f = \frac{1}{\cos \theta_i + \cos \theta_r} \cdot R_B(\beta, g) \cdot S(\beta).
\]  

(10)

The first factor is the well-known Lommel-Seeliger scattering term arising from the total distance traveled through the attenuating medium in the course of reflecting from a typical particle. The second term is called the retroreflection function and decreases monotonically in value from 2 for \( 2\beta = 0 \) (backscattering) to 1 for \( 2\beta \geq \pi/2 \). It depends on a constant, \( g \), of the order of unity, which is called the compaction parameter because it is related to the fractional volume of the medium occupied by the particles. The last factor is the scatter function, defined as the ratio of the scattered power per unit solid angle to the power incident on a particle by a plane wave, averaged over all orientations and shapes of the particles. Taking the incident plane wave to be of unit intensity, this implies

\[
S(\beta) \equiv \left\{ \frac{dP_r}{A_\perp d\Omega_r} \right\}
\]

(11)

where \( A_\perp \) is the cross-sectional area of the particle (i.e., the projected area of the particle onto an incident wave front). Notice that \( S(\beta) \) is approximately equal to the ensemble average of the cosine-corrected BRDF (Stover, 1990) of the particles, \( dP_r/\sin(\theta_i) \); the equality becomes exact for pancake-shaped (i.e., planar) particles. The scatter function is written by Hapke as \( S(\beta) = \rho \Sigma(\beta) \), where \( \Sigma(\beta) \) has been normalized such that

\[
4\pi \int_0^{\pi/2} \Sigma(\beta) \sin(2\beta) \, d\beta = 1.
\]

(12)

Both \( S(\beta) \) and \( \Sigma(\beta) \), just like \( f \), have units of sr\(^{-1}\). For example, \( S(\beta) = 1/4\pi \) describes isotropic scattering from smooth, infinitely conducting (i.e., ideally specular) spheroids—if the conductivity is not infinite, then Eq. (6) implies that there will be a broad forward scattering peak (Van de Hulst, 1957) since the Fresnel reflectances only become equal to unity at grazing angles—while \( \Sigma(\beta) = 2\sin(2\beta) + (\pi-2\beta)\cos(2\beta)/3\pi^2 \) describes backscattering from Lambertian spheres, as first calculated by Schönberg. As can be seen from Hapke Fig. 6, substituting this latter scattering function into Eq. (10) results in a strong backscattering peak. This is also the explanation for the glory seen when looking down on clouds from an airplane with the sun at
one’s back and for the hot spots which appear on vegetation in remote sensing.

Experimental retroreflectance measurements of MgCO$_3$, BaSO$_4$, sulfur, and white, red, blue, and black Nextel paints were later performed by Egan and Hilgeman (1976) using a cube beamsplitter, which was cleverly followed (rather than preceded) by a chopper so as to reject light scattered by the room or the prism. Both a bandpass-filtered tungsten iodide lamp and a 632.8-nm HeNe laser were used as sources. In all cases an opposition effect was observed that could reasonably be described by Eq. (10), except for the blue and black paints under laser illumination which exhibited an anomalously strong backscattering peak attributed to interference effects. Presciently, the opposition effect of the moon is now known to result not from shadow hiding but instead from coherent backscatter (Hapke et al., 1993). This phenomenon, related to weak localization, arises as follows (Wolf and Maret, 1985). Suppose that a wave of incident propagation vector $k_0$ experiences $m$ elastic scattering events, where $m \geq 2$. Let $k_n$ denote the propagation vector after the $n$th event, so that $k_n$ points in the direction of observation. In the case of backscattering, $k_n = -k_0$, and hence some of the incident field can also follow the time-reversed path $-k_n \to -k_{n-1} \to \ldots \to -k_0$. The phase difference between these two paths is obviously zero and thus they will interfere constructively, giving double the signal enhancement effect is responsible for lunar backscattering was provided by examining the circular polarization ratio of scattered radiation from soil samples with incident circular laser light. Shadow hiding involves primarily single scattering so that the helicity should be reversed, while coherent backscatter involves multiple scatterings, many of which are into the forward direction, and hence the original polarization should be partially preserved, and this is what is in fact observed. The angular width of this coherent backscatter peak is supposed to approximately equal $\lambda l / l$, where $l$ is the transport mean free path for photons in the medium; for strongly absorbing particles as in the case of the moon, $l$ is roughly equal to the average spacing between scatterers. Analysis of the data implies $l = 1 \, \mu m$, in contrast to the 40 $\mu m$ mean particle size in the lunar dust; hence, the scatterers must be small asperities on the grain surfaces, rather than the particles themselves. This is consistent with the observation of coherent backscatter from wavelength-sized roughness on well-characterized surfaces (O’Donnell and Mendez, 1987).

Returning to Oren and Nayar’s model, suppose that a V-cavity is illuminated from the right. Then the left facet will be brighter than the right one because it receives more light. Viewed from the left, an observed sees principally the darker right facet and comparatively little of the foreshortened left facet. But as he moves toward the source direction, the fraction of the brighter area increases while that of the darker decreases and hence the BRDF increases in the backscattering direction and is inherently non-Lambertian. Specifically, the BRDF for a single Lambertian facet whose normal is inclined at polar angle $\theta_n$ and azimuthal angle $\phi_n$ relative to the macroscopic surface normal is

$$f = \rho \cos \theta_n \left[1 + \tan \theta_i \tan \theta_n \cos(\phi_i - \phi_n)\right] \left[1 + \tan \theta_i \tan \theta_n \cos(\phi - \phi_n)\right],$$

where $\rho$ is the albedo of the facet. This expression is then multiplied by a geometric attenuation factor $G(\hat{i}, \hat{n}, \hat{F})$ to account for shadowing and masking, and the result is averaged over $\phi_n$, assuming a uniform distribution of azimuthal orientations of the V-grooves. Finally, that is multiplied by the density function of facet normals $\Xi(\theta_n)$ and integrated with respect to $\sin \theta_n d \theta_n$ over the hemisphere. Oren and Nayar chose $\Xi(\theta_n) \cos \theta_n$ to be a Gaussian with zero mean, which differs slightly from Torrance and Sparrow (1967) or Maxwell and Weiner (1974) who chose $\Xi(\theta_n)$ itself to be such. Finally, two-bounce interreflections were taken into account by integrating over all positions on the two faces of a V-groove which connect directions $\hat{i}$ and $\hat{r}$ geometrically, where shadowing and masking determine the limits of integration, and finally again integrating over $\phi_n$ and $\theta_n$. The singly scattered and doubly scattered BRDFs were then added to give the total BRDF, which exhibits reciprocity. Some of the required integrals are fairly complicated and were evaluated numerically or approximated functionally. The resulting
BRDF is nearly Lambertian for small angles of incidence, wherein both facets of every V-cavity have similar irradiance, as well as for azimuthal angles 90° out of the plane of incidence since the relative irradiance of both facets is then approximately constant with respect to the polar angle of reflection. But there is a strong backscattering peak when the source and viewing directions coincide, as well as an interreflection enhancement in the forward direction, and the rendered image of a curved surface illuminated from the viewer direction can be made very flat, mimicking the appearance of the moon or of a photographed clay vase for example. Of course the model reduces to pure Lambertian scattering if the roughness (as parametrized by the Gaussian standard deviation in the facet density function) is set equal to zero.

The facets in Torrance and Sparrow’s or Oren and Nayar’s model are assumed to be large compared to the wavelength λ of the incident light. More specifically, any non-composite, isotropic surface (or any individual component of a composite surface) can be characterized by two roughness scales—an out-of-plane length quantified by the rms surface height variation σ and an in-plane distance parametrized by the correlation length l. In the facet models, both of these parameters (for the coarse component in Oren and Nayar’s case) must be much larger than λ. Mathematically (Bennett and Porteus, 1961), the surface height profile is taken to be \( z = \zeta(x,y) \) with the zero level set equal to the mean,

\[
\langle z \rangle = \lim_{X,Y \to \infty} \frac{1}{XY} \int_{-Y/2}^{Y/2} \int_{-X/2}^{X/2} \zeta(x,y) \, dx \, dy = 0,
\]

where the sample has been taken to be infinite in area, macroscopically spanning the xy-plane, for simplicity. With these definitions, the mean square surface height is given by the variance,

\[
\sigma^2 \equiv \langle z^2 \rangle = \lim_{X,Y \to \infty} \frac{1}{XY} \int_{-Y/2}^{Y/2} \int_{-X/2}^{X/2} \zeta^2(x,y) \, dx \, dy,
\]

and the autocorrelation (or autocovariance) function is (Hecht, 1998)

\[
A(s,t) \equiv \langle \zeta(x,y)\zeta(x-s,y-t) \rangle = \lim_{X,Y \to \infty} \frac{1}{XY} \int_{-Y/2}^{Y/2} \int_{-X/2}^{X/2} \zeta(x,y)\zeta(x-s,y-t) \, dx \, dy,
\]

so that \( A(0,0) = \sigma^2 \). For an isotropic surface, \( A(s,t) = A(t,s) \) so that the correlation function is only a function of a single variable \( \tau \equiv \sqrt{s^2 + t^2} \) known as the lag; the correlation length \( l \) is defined as that lag for which \( A \) diminishes to 1/e of its peak value (i.e., to \( \sigma^2/e \)). There are two commonly used statistical distributions for rough surfaces (Barrick, 1970). The first assumes Gaussian correlations,

\[
A(s,t) = \sigma^2 e^{-\left((s^2+t^2)/l^2\right)},
\]

for which the rms slope of the surface can be shown to be

\[
m = \frac{\sqrt{\left(\frac{\partial \zeta}{\partial x}\right)^2 + \left(\frac{\partial \zeta}{\partial y}\right)^2}}{l} = \frac{2\sigma}{l},
\]

(which equals the rms value of tanθ in the facet models and hence is determined in turn by the density function). The other supposes the surface height autocorrelation function to be exponential,

\[
A(s,t) = \sigma^2 e^{-\sqrt{s^2+t^2}/l},
\]

in which case \( m \) turns out to be undefined, because such surfaces are jagged with many vertical facets, describing, as an example in remote sensing, an urban area including buildings.

The requirement that \( \sigma \gg \lambda \), as well as the naive model of V-grooves having coplanar top edges, can be relaxed by invoking wave rather than geometrical optics. Nevertheless, many simplifying assumptions remain in order to make the physics tractable; the general problem of
optical scattering by an arbitrarily specified surface has not been solved, even without considering volumetric scattering of the portion of the beam transmitted below the first surface. Briefly, the goal is to find the reflected electromagnetic field for some known incident field and characterized surface by satisfying the required boundary conditions. Typically shadowing, masking, multiple reflections, and surface waves are neglected. One approach has been developed in great detail by the radar scattering community (Beckmann and Spizzichino, 1987); note that the (differential) radar cross section \( \sigma \) (having units of area) is related to the BRDF by \( f = \sigma / 4 \pi A \cos \theta_1 \cos \theta_r \), where \( A \) is the (macroscopic) illuminated surface area of the sample (Ruck et al., 1970). O’Donnell and Mendez (1987) have verified that this theory is in good agreement with optical measurements on appropriately fabricated samples.

In detail, Maxwell’s equations imply that the electric field satisfies a wave equation, and if the time dependence is separated out (Fourier analyzing the constituent frequencies if necessary), the Helmholtz equation is obtained for any scalar component of the field (typically the incident and reflected radiation are decomposed into \( s \) and \( p \) polarizations). Green’s theorem can be used to recast this differential equation as the Kirchhoff integral, which expresses the field anywhere in space in terms of the electric field and its gradient at every point on any closed surface enclosing the spatial point of interest. In the present application, the surface is taken to be that of the object together with an enclosing hemisphere at infinity, with the latter giving zero contribution to the result. In accord with Huygens’ principle, each point on the material surface is taken to be the source of a spherical wave, and the Fraunhofer far-field limit is considered, in which the incident and scattered beams are taken to be plane waves. Finally, the Kirchhoff approximation is invoked to obtain the electric field and its normal derivative on the reflecting surface, by assuming that the scattering from any point on the surface is described by the Fresnel amplitude reflection coefficients, \( r \) in Eq. (6), from a smooth plane tangent to that point. This is the most serious restriction of the validity of this approach, as it obviously requires that the radii of curvature of any surface irregularities be large compared to the wavelength, or equivalently, that \( l >> \lambda \), implying a gently rolling surface. A formal integral solution to the problem of scattering by a rough surface of finite conductivity is then obtained. Unfortunately it is too complicated to evaluate for real situations of interest unless the relevant amplitude reflection coefficient, which depends on the angle of incidence relative to the local surface normal, is constant across the surface. In practice, this occurs either because the conductivity can be taken to be infinite or because the surface roughness is sufficiently small that we can replace the reflection coefficient by its average value. For the simpler, perfectly conducting case, integration by parts gives the scattered field in terms of a Fourier transform over the area \( A \) of the isotropic surface,

\[
\frac{E_r}{E_{r\text{,smooth}}} = F(\theta_1, \phi_r, \phi_r - \phi_i) \int_{A} e^{i(k_i \cdot -k_r) \cdot r_s} dx dy
\]

(20)

neglecting edge effects, where \( k_i \) and \( k_r \) are the incident and reflected propagation vectors, respectively, with magnitude \( k \equiv 2 \pi / \lambda \), \( E_{r\text{,smooth}} \) is the field which would be diffracted into the specular direction if the finite-sized, infinitely-conducting surface were smooth and the incident wave were \( s \)-polarized, \( F = [1 + \cos \theta_1 \cos \theta_r + \sin \theta_1 \sin \theta_r \cos (\phi_r - \phi_i)]/\{\cos \theta_r (\cos \theta_1 + \cos \theta_r)\} \), and \( r_s = (x, y, z) \) is a point on the surface. Accounting for the Huygens-Fresnel Principle (cf. Sec. 10.3.1 of Hecht, 1998), the projected area of the sample \( (A \cos \theta) \), and the solid angle subtended at the detector \( (d\Omega) \), I find that the BRDF is given by

\[
f = \frac{A \cos \theta_1}{\lambda^2 \cos \theta_r} \left( \frac{E_r}{E_{r\text{,smooth}}} \right)^2,
\]

(21)

which satisfies Helmholtz reciprocity. For a statistically rough surface, Eq. (21) needs to be averaged by multiplying it by the normalized height distribution \( p(\zeta) \) and integrating over all \( \zeta \). The usual assumption is that this distribution is Gaussian, i.e.,
\[ p(\zeta) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\zeta^2/2\sigma^2}. \]  
(22)

Notice that Eq. (21) will involve a quadruple integral (over say \(dA'dA\)) which can be recast in terms of the autocorrelation function. Assuming that this is also Gaussian and given by Eq. (17), a rather complicated final expression for the scattering results.

Two special cases are of most interest however. For a slightly rough surface \((\sigma \ll \lambda)\), the scattering is predominantly specular and the directional-conical reflectance is approximately

\[
\rho_{\text{spec}}(\theta_i, \phi_i; \Delta \Omega_{\text{spec}}) = e^{-(2k\sigma \cos \theta_i)^2},
\]
(23)

where \(\Delta \Omega_{\text{spec}}\) spans the specular lobe centered about \(\theta = \theta_i\) and \(\phi = \phi_i + \pi\) and has a projected value of roughly \(\lambda^2/A\) due to diffraction from the finite-sized surface \(A\) (which is determined, for example, by the usual product of sinc-squared functions for a rectangular surface). Equation (23) is a well-known result nicely derived from the Fraunhofer diffraction formula by Davies (1954). Bennett and Porteus (1961) have experimentally verified that it gives the specular reflectance, at sufficiently long wavelengths, of a slightly rough surface relative to a smooth surface of the same material. Good agreement was obtained with measurements in the mid-infrared for glass disks roughened with \(-10\mu m\) grit, overcoated with aluminum, and characterized by stylus profilometry. Normal incidence was employed, thus minimizing effects such as shadowing and any polarization dependence. These measurements can also be described in terms of the total integrated scatter (TIS), defined as the ratio of the diffusely to the specularly scattered powers. Writing the specularly, diffusely, and total scattered powers as \(P_{\text{spec}}, P_{\text{diff}},\) and \(P_{\text{tot}} = P_{\text{spec}} + P_{\text{diff}}\), respectively, one sees for a slightly rough surface that

\[
\text{TIS} = \frac{P_{\text{tot}} - P_{\text{spec}}}{P_{\text{spec}}} = \frac{1 - \rho_{\text{spec}}}{\rho_{\text{spec}}} \approx (2k\sigma \cos \theta_i)^2,
\]
(24)

where the second equality presumes that slight roughening merely redistributes the total power scattered by a smooth sample without diminishing it, as follows from the assumption of infinite conductivity. Note that to the level of the approximations used above, the TIS is also equal to

\[ 1 - \left( P_{\text{spec}} / P_{\text{tot}} \right) = 1 - \left( P_{\text{spec}} / P_{\text{spec,smooth}} \right), \]

where, recalling Eq. (5), \(P_{\text{spec,smooth}} = R(\theta)P_i\) is the specularly reflected power by a smooth surface of the same material, thus explaining other definitions of the TIS used in the literature (e.g., Elson and Bennett, 1979b). Church et al. (1977) have shown, from the Rayleigh-Rice perturbation theory discussed below, that Eq. (24) holds regardless of the specific form chosen for the height distribution function \(p(\zeta)\) in Eq. (22) and hence is not dependent on the assumption of Gaussian statistics—see, for example, the discussion in connection with Eq. (34) below. A particularly nice feature of TIS measurements is that no reference sample is necessary: the diffusely reflected power can be measured using an integrating sphere and one detector, while a small hole (coincident with the input hole for normal incidence) and a second detector is used to measure the specularly reflected power (Stover, 1990). Also note that Eq. (24) provides a quantitative basis for the Rayleigh criterion for the roughness of a sample, wherein a surface is respectively considered rough or smooth if \(4\pi\sigma \cos \theta_i/\lambda\), which equals the phase difference between two specular rays reflected from points separated by a height \(\sigma\) on the surface, is large or small compared to \(\pi/2\) or approximately 1 (Beckmann and Spizzichino, 1987).

Equation (23), giving the specular reflectance of a slightly rough surface, is in fact proportional to \(<|E_r|^2\>\). However, the total scattered intensity is more exactly given by \(<|E|^2\> = <|E_r|^2\> + <|E_d - <E_r>|^2\>\). The angular distribution of the diffusely scattered radiation is determined by the second term on the right-hand side of this equality, which defines the variance of the complex scalar amplitude \(E_d\), given in turn by the sum of the variances of its real and imaginary parts. In general, this depends upon the joint probability distribution of finding the height to be \(\zeta\), at a lag \(\tau\) away from a point where the height is \(\zeta_1\). However, for a slightly rough surface the integral of interest can be expanded to first order in \(\zeta/\lambda\) and \(\sigma/\lambda\), with the result that
the variance depends only on the autocorrelation function of Eq. (17). The diffuse BRDF is then approximately found to be (Beckmann and Spizzichino, 1987)

\[
f_{\text{diff}} = \frac{k^4 l^2 \sigma^2}{4\pi} \left[ 1 + \cos \theta_i \cos \theta_r + \sin \theta_i \sin \theta_r \cos (\phi_r - \phi_i) \right]^2 \cos \theta_i \cos \theta_r \\
\exp \left[ -k^2 l^2 \left\{ \sin^2 \theta_i + \sin^2 \theta_r + 2 \sin \theta_i \sin \theta_r \cos (\phi_r - \phi_i) \right\} / 4 \right].
\]

(25)

Similar results have been obtained by Davies (1954) and Leader (1979) except for the angular prefactor—as discussed by Stover (1990), as well as in a footnote on page 75 and in Appendix A of Beckmann and Spizzichino (1987), the form of this angular term depends on the obliquity factor used, of which considerable variation exists in the literature.

The other important case of Eq. (21) averaged using Eq. (22) is very rough scattering (\(\sigma \gg \lambda\)), for which the reflection is predominantly diffuse and given by the BRDF

\[
f = \frac{1}{\pi m^2} \left[ 1 + \cos \theta_i \cos \theta_r + \sin \theta_i \sin \theta_r \cos (\phi_r - \phi_i) \right]^2 \cos \theta_i \cos \theta_r (\cos \theta_i + \cos \theta_r)^4 \\
\exp \left[ \frac{\sin^2 \theta_i + \sin^2 \theta_r + 2 \sin \theta_i \sin \theta_r \cos (\phi_r - \phi_i)}{m^2 (\cos \theta_i + \cos \theta_r)^2} \right]
\]

(26)

where, as defined in Eq. (18), \(m\) is the rms surface slope. In Davies (1954), the argument of the exponential is larger by a factor of 2, apparently stemming from a different assumed form for the surface-height joint probability distribution.

It is emphasized again that the tangent-plane approximation inherent in the above physical-optics method implies that the solution will only be correct in the high-optical-frequency limit (\(\lambda \ll l\)) and in this respect is not more accurate than geometrical optics, although it can be applied to a wider variety of target geometries. A perturbation approach can be used to approximate the scattering at lower frequencies (\(\lambda < l\); obviously at very low frequencies a rough surface will ultimately appear smooth and scatter specularly) and has the added advantage that it is explicitly vector based. However, the relevant perturbation is of the surface height and hence the method only applies to slightly rough samples (\(\sigma \ll \lambda\)). The basic idea was first proposed by Rayleigh in 1895 and later extended by Rice (1951), so that it is today known as the Rayleigh-Rice method. The height profile \(\zeta(x,y)\) is Fourier analyzed into its sinusoidal components \(\hat{Z}(k_x, k_y)\). Any individual component acts like a periodic grating, scattering an incident plane wave into a set of directions given by the 2D grating equations. Assuming the surface is only slightly rough, the diffuse scattering will be dominated by the first-order peaks, so that

\[
(k_r - k_i) \cdot \hat{x} = k \left( \sin \theta_r \cos \phi_r - \sin \theta_i \right) = k_x
\]

\[
(k_r - k_i) \cdot \hat{y} = k \sin \theta_r \sin \phi_r = k_y
\]

(27)

where \(k_i\) and \(k_r\) can be either positive or negative so as to include all four first-order peaks, and where I have followed the usual convention (e.g., Beckmann and Spizzichino, 1987; Church and Zavada, 1975) of orienting the \(x\)-axis so that \(\phi = -\pi\). Equation (27) can be viewed as an expression of conservation of linear momentum. The total electric field above the surface (which for simplicity is taken to be infinitely conducting) in the case of incident \(s\)-polarization (which avoids coupling to surface waves) is then

\[
E = E_0 e^{ik_r \cdot r} - E_0 e^{ik_i \cdot (0,0) \cdot r} - \sum_{k_x, k_y} E_{0r}(k_x, k_y) e^{i k_r \cdot (k_x, k_y) \cdot r},
\]

(28)

where the 0 subscripts on the electric fields denote the complex amplitudes, suppressing the unimportant time dependences. The first term on the right-hand side of Eq. (28) is the incident field; the second is the specularly reflected field, whose amplitude is equal to that of the incident
field to lowest order in the roughness and whose sign arises from the 180° phase shift for external TE reflection from Eq. (6a); and the last term is the diffusely reflected field, given by the sum of the first-order diffraction peaks. (Note that Rice writes $k_x = mK$ and $k_y = nK$, where $K$ is some arbitrarily small fundamental spatial frequency, which in practice can be chosen by requiring $2\pi K$ to be of the order of the largest experimentally observable spatial wavelength, namely the sample or beam diameter. Furthermore, there are upper limits on the values of $k_x$ and $k_y$, beyond which the diffracted orders disappear into the surface; for example, one must have $(k_x^2 + k_y^2)^{1/2} \leq k$ at normal incidence, i.e., $\lambda < l$. Thus, the above summation can actually be considered to be over a finite set of integers $m$ and $n$.) Maxwell’s equation $\nabla \cdot E = 0$ requires that each component wave in the scattered field be transverse, $k \times E_0 = 0$. Furthermore, electromagnetic boundary conditions must be imposed at the surface. For example, the two tangential components of the total electric field must vanish on the surface, $\hat{n} \times E|_{z=\zeta} = 0$, where $\hat{n}$ is the local surface normal given by $\hat{n} = \frac{(-\partial \zeta / \partial x, -\partial \zeta / \partial y, 1)}{\sqrt{1 + (\partial \zeta / \partial x)^2 + (\partial \zeta / \partial y)^2}}$, or approximately $(-\partial \zeta / \partial x, -\partial \zeta / \partial y, 1)$ using the perturbation assumption (which implies the surface slopes are gentle, again consistent with $\lambda < l$). These first derivatives of the surface profile are given by the Fourier transforms of $k_x Z(k_x, k_y)$ and $k_y Z(k_x, k_y)$. A matrix equation is thus obtained for the Cartesian components of the unknown amplitudes $E_0$, which (as for $\hat{n}$) are expanded to lowest nonzero order in $\zeta$. The statistical character of the surface profile is now introduced by assuming that $Z(k_x, k_y)$ is distributed normally about zero and that the variables $k_x$ and $k_y$ are independent. In particular, the ensemble average of the squared modulus of $Z(k_x, k_y)$ defines the power spectral density (PSD),

$$W(k_x, k_y) = \frac{1}{A} \left[ \left\langle \int_{A} \zeta(x,y) e^{i(k_x x + k_y y)} \, dx \, dy \right\rangle \right]^2,$$

where $A$ denotes the area of the sample. (In the present context, the $1/A$ prefactor normalizes the scattered power to the incident power.) According to the Wiener-Khinchin relation, Eq. (30) can also be expressed as the Fourier transform of the autocorrelation function. In polar coordinates, the azimuthal integration yields a Bessel function of order zero for an isotropic surface (Elson and Bennett, 1979b), so that $W$ is only a function of $(k_x^2 + k_y^2)^{1/2}$. Specifically, assuming a Gaussian autocorrelation function as given by Eq. (17), the radial integration can also be performed, giving

$$W(k_x, k_y) = \pi l^2 \sigma^2 e^{-(k_x^2 + k_y^2)l^2/4},$$

while for the exponential surface height correlations of Eq. (19), one obtains

$$W(k_x, k_y) = \frac{2\pi l^2 \sigma^2}{\left[ 1 + (k_x^2 + k_y^2)l^2 \right]^{3/2}}$$

which becomes a Lorentzian in the case of 1D roughness, $\zeta(x)$. Note that, strictly speaking, Eq. (32) cannot be used in the present context because it implies (infinitely) steep surface slopes, as discussed in connection with Eq. (19); however, many researchers have ignored this point and used it anyway, implicitly assuming that the correlation function eventually becomes non-exponential for lags below the range of measurement (which is on the order of $\lambda$ in optical experiments and of the radius of the stylus when using mechanical profilometry). Specifically, the autocorrelation function must have zero slope at zero lag to be physically meaningful (Elson and Bennett, 1979a). Elson (1975; also see Elson and Ritchie, 1974) has performed an analysis which is essentially equivalent to the preceding Rayleigh-Rice vector theory, except that he has used the full
incident light is polarized but neither particular, the cross-polarization terms incident beam and can be considered a generalization of the Fresnel reflectances of Eq. (6). In otherwise be masked by the microroughness surface scattering. such as those due to surface particles or bulk defects (Germer, 1997), whose signals might vibration in general. Both of these effects can be used to probe other scattering mechanisms, (lateral) scattered light remains strongly linearly polarized, albeit with a rotated plane of Furthermore, ellipsometry indicates that, for incident linearly polarized light, the out-of-plane angles (e.g., pp). An interesting feature of the data is that the small-angle scattering [which by Eq. (27) implies \( \sin(\theta_i) \approx \sin(\theta_r) \), i.e., a low-spatial-frequency grating and an angle of incidence not too far off-normal], integrating it with respect to \( \cos(\theta_r) d\Omega_r \), over the outgoing hemisphere (implicitly excluding the specular beam), and dividing that by \( R(\theta) \) gives the TIS, which agrees with Eq. (24). This result can also be written, to lowest order in \( \sigma \), as TIS = 2\( p \), where \( p \) is called the diffraction efficiency and is defined as the ratio of the power diffracted into the +1 and 0 orders (Stover, 1975; Church et al., 1977), with the factor of 2 arising from the fact that the same amount of power is diffracted into the +1 and –1 orders in the small-angle-scattering limit. Equation (33) has been compared to experimental measurements (Germer et al., 1997) on a fabricated silicon microroughness standard having a pseudorandom distribution of circular pits (with ~1-\( \mu \)m diameters and 1-nm depths) using a cw doubled Nd:YAG laser. Excellent agreement resulted for an incident angle of \( \theta_i = 45^\circ \), reflected polar angles of \( \theta_r = 30^\circ, 45^\circ \), and 60° and azimuthal angles of \( \phi_r = 10^\circ–170^\circ \), and the four polarization combinations ss, sp, ps, and pp. An interesting feature of the data is that the pp scattering vanishes for certain out-of-plane angles (e.g., \( \phi_r = 60^\circ \) for \( \theta_r = 45^\circ \)) which are the bidirectional analogs of the Brewster angle. Furthermore, ellipsometry indicates that, for incident linearly polarized light, the out-of-plane (lateral) scattered light remains strongly linearly polarized, albeit with a rotated plane of vibration in general. Both of these effects can be used to probe other scattering mechanisms, such as those due to surface particles or bulk defects (Germer, 1997), whose signals might otherwise be masked by the microroughness surface scattering.

The optical factor \( \mathcal{Q} \) describes the polarization of the reflected beam relative to that of the incident beam and can be considered a generalization of the Fresnel reflectances of Eq. (6). In particular, the cross-polarization terms \( Q_{xy} \) and \( Q_{yx} \) are zero for scattering into the plane of incidence. For the common case of in-plane (longitudinal) ss scattering, it is convenient to note (Stover, 1990) that the optical factor is simply the geometric mean of \( R_x(\theta) \) and \( R_y(\theta) \). If the incident light is polarized but neither s nor p (i.e., linearly polarized at an oblique angle,
circularly polarized, or elliptically polarized), then for both the specular and diffuse components even the singly scattered light is depolarized. To put this on a more precise footing, the polarization factor $p$ of a wave is defined as $E_0^0/E_{0i}$, where the complex amplitude $E_0$ has been resolved into its vertical and horizontal components (Beckmann and Spizzichino, 1987). I choose to follow the sign convention wherein any scalar component of the field is written as

$$E = |E_0|^e^{i(k \cdot r - \omega t + \phi)} = E_0 e^{i(k \cdot r - \omega t)},$$

where $\omega = ck$ in vacuum, while some authors choose the opposite signs for $k \cdot r$ and $\omega t$. Accordingly, $\text{Im}(p)$ implies linear polarization (e.g., $p = 0$ is horizontal and $p = \infty$ is vertical), while $\text{Im}(p)$ greater or less than 0 refers to left-handed or right-handed polarization, respectively (e.g., $p = i$ is left circular and $p = -i$ is right circular). If the polarization of the incident wave is described by $p_i$ and that of the reflected wave by $p_r$, then the (complex) depolarization factor $q$ is defined by the relation $p_r = qp_i$; specifically, $q = 1$ implies no depolarization, while $q = 0$ or $\infty$ corresponds to a filter, scattering only one component of the incident light, as occurs at the Brewster angle for example. [Note carefully that depolarization here refers to a change in the state of polarization of a beam, not in its degree of polarization. Both the incident and scattered waves are assumed to be fully polarized. However, many workers (e.g., Renau et al., 1967) use the term to refer instead to an increase in the randomly polarized fraction of a beam.] For an ideal specular reflector, $q = -1$, so that an infinite, flat, perfect conductor will depolarize all incident polarizations except $s$ and $p$, changing both the orientation of the polarization ellipse and its handedness; this same result holds for longitudinal scattering by a rough surface (of infinite conductivity) if the tangent-plane approximation holds (i.e., $l >> \lambda$).

If a smooth plane has finite conductivity instead, then $q = q_\sigma$ (e.g., Jenkins and White, 1957); in contrast, in the case of a dielectric remains linearly polarized but with a rotated plane of vibration, and in the case of a metal is elliptically polarized (Jenkins and White, 1957); in contrast, $q = -1$ at normal incidence and +1 at grazing incidence.

In the case of lateral scattering, even incident $s$ or $p$ light is depolarized. Consider, for example, a vertically polarized wave at near grazing incidence scattering off a perfectly conducting surface element tilted upwards about the $x$ axis by $45^\circ$. Resolving the incident field into tangential and normal components and applying the boundary conditions, the reflected wave is found to be horizontally polarized. This effect is responsible for the depolarization of television waves by tin rooves in cities. The general result for specular reflection off facets of finite conductivity is shown from spherical geometry (Beckmann and Spizzichino, 1987) to be

$$p_r = \frac{p_i(r_i \tan \beta_1 \tan \beta_2 + r_p) + (r_p - r_i) \tan \beta_1}{r_s + r_p \tan \beta_1 \tan \beta_2 - p_i(r_s \tan \beta_2 - r_p \tan \beta_1)},$$

where

$$\beta_1 \equiv \sin^{-1}\left[\frac{\sin \theta_r \sin \phi_r}{\sqrt{1 - (\cos \theta_i \cos \theta_r - \sin \theta_i \sin \theta_r \cos \phi_r)^2}}\right]$$

and

$$\beta_2 = \cos^{-1}\left[\cos \beta_1 \cos \phi_r - \sin \beta_1 \cos \theta_r \sin \phi_r\right].$$

with $\phi = -\pi$, as in Eq. (27). It is easy to check that for longitudinal scattering ($\phi = 0$), one gets $\beta_1 = \beta_2 = 0$, so that $q = r_i/r_s$, as above. These expressions are already rather complicated, without even considering the depolarization of waves diffracted into non-specular directions, as occurs in the Rayleigh-Rice model for example. Note that knowledge of the optical factors is not sufficient to describe scattering from and to states of arbitrary polarization, because the four $Q$ values involve intensity rather than amplitude ratios and hence the necessary phase information is not available. For example, a beam linearly polarized at $45^\circ$ to the vertical and a circularly polarized
wave both have equal-intensity $s$ and $p$ components and hence cannot be distinguished on that basis alone. Unfortunately, no general analytical solution to the problem of depolarization by a rough surface exists.

To circumvent this situation and the limitations inherent in the Kirchhoff method (i.e., the tangent-plane approximation $l \gg \lambda$ and the assumption of infinite conductivity) or in the Rayleigh-Rice theory (namely, the assumption of slight roughness $\sigma \ll \lambda$ and gentle slopes $\lambda \ll l$), direct numerical solutions of the electromagnetic equations have been undertaken. In essence, one starts from a vector form of the Kirchhoff integral (Jackson, 1975)—also known as the Chu-Straton integrals (Ruck et al., 1970)—for the electric and magnetic fields, $E$ and $H$, which express the total fields (or more simply, the scattered fields if the incident fields are subtracted off) at any point in space in terms of their values on the sample surface and a free-space Green function. In turn, the surface fields can be related to the induced surface current density $J$. Finally, $J$ can be obtained from a similar integral by relating the scattered fields on the surface to the incident fields via the boundary conditions, taking care to handle the singularity when the argument of the Green function is zero (i.e., when the source point is identified with the field point). These integral equations can also be obtained more directly from the optical extinction theorem (Soto-Crespo and Nieto-Vesperinas, 1989). To solve them numerically, the finite surface is chopped into a discrete number of sampling points, say $N$ in all. The elements of surface area, $\text{sec}\theta_x dxdy$, and the surface-normal unit vectors $\hat{n}$ (where $\hat{n} \cdot \hat{z} \equiv \cos\theta_x$) are determined by $\zeta(x,y)$ from Eq. (29). The integrals expressing the boundary conditions can then be written as two $N \times N$ matrix equations for the $N$ values of $E$, and $H$, on the surface in terms of the $N$ values of the two tangential components of $J$. These equations are inverted to give $J$, which is inserted in turn into the finite approximations for the integrals for $E$, and $H$, completing the solution. This is known as the point-matching method. The $N$ sampled values of the surface roughness profile are generated by a Monte Carlo procedure constructed to give zero mean with the desired root variance $\sigma$ and autocorrelation function $A(\tau)$. The scattered power is proportional to $\langle |E|^2 \rangle$, where the average is over an ensemble of such sequences of sampled values.

For example, Soto-Crespo and Nieto-Vesperinas (1989) have computed the scattering from 1D infinitely-conducting random (Gaussian-correlated) or periodic surfaces. For this geometry, the Green function is a zeroth-order Hankel function of the first kind. Interestingly, there is an intermediate range of roughness (between the regimes of mostly specular scattering and enhanced backscattering at small and large values of $\sigma\lambda$, respectively) for which the reflection is very nearly Lambertian, at least for near-normal angles of incidence. Furthermore, these researchers compared their numerical results to the analytical predictions of the Kirchhoff approximation and were thus able to prepare a useful graph indicating the range of values of $\sigma l$ versus $l/\lambda$ for given angles of incidence for which the two integrated scattered powers agree with each other to within a specified percentage. Saillard and Maystre (1990) have considered 1D random surfaces of finite but large conductivity, applicable for instance to gold at infrared (IR) wavelengths. In this case, the kernels of the integral equations are sharply peaked and hence difficult to accurately evaluate by the above numerical procedures. So instead, an “impedance boundary condition” is imposed, whereby a local linear relationship between the total field in the metal and its normal derivative on the surface is assumed. Knotts et al. (1993) have numerically evaluated the complete Mueller matrix for these same conditions and compared the results to experimental measurements on well-defined samples, prepared by overcoating a photoresist plate exposed to a suitable 1D laser pattern. Qualitative agreement resulted when Gaussian statistics were employed, but detailed agreement required that the sequences of surface-height values be taken directly from digitized profilometer data rather than be generated by the Monte Carlo technique, indicating that rather subtle statistical properties of a surface can play a significant role in determining the scattering. This emphasizes the most important limitation of such numerical techniques: they must be repeated from scratch for each specific set of sample properties and illumination characteristics.
In this section, an assortment of experimental papers are briefly discussed, but no attempt at completeness is made—for this purpose, reviews such as the one by Asmail (1991) should be consulted. It is logical to begin with the reflectometers used to measure BRDFs in the laboratory. Typical examples are described by Roche and Pelletier (1984), Zaworski et al. (1993), and Sandmeier et al. (1997). At their hearts are a goniometer which permits one to choose values of $\theta_i$, $\phi_i$, $\theta_r$, and/or $\phi_r$ in certain ranges, to within specified angular resolutions and source and detector solid angles. Two common design geometries are: type (i) where the source and detector are each mounted on a pair of arc rails so that they can be scanned throughout the hemisphere above a stationary sample; and type (ii) for which the source is fixed in place, the sample rotates about the $x$ and $y$ axes defining its surface, and the detector orbits around the sample in a single plane. Various choices for the collimation or focusing of the incident light are discussed in Nicodemus et al. (1977). Sometimes the sample is rapidly spun about its normal direction to average out inhomogeneities such as laser speckle, and the incident beam chopped so that lock-in detection can be used to eliminate scattering by ambient light. Oppenheim et al. (1994) have made measurements on various recommended standards of diffuse (nearly Lambertian) reflectance in the IR used in calibrating such instruments.

Moving on to applications in the domains of interest outlined in the introduction, Snyder and Wan (1996) have used a Fourier-transform IR spectrometer mounted on a type (i) goniometer to investigate soil samples. It proved necessary to make measurements both with the source on and off, in order to correct for sample heating by it. The ultimate goal is to determine the BRDF for analysis of satellite-based thermometric imaging of earth’s surface. Stavridi et al. (1997) have examined brick, tile, and concrete with a CCD camera in a type (ii) geometry for architectural applications. As is typical of many experimental papers, these researchers have fitted their data by combining several models discussed in the previous section. Specifically, they chose a linear combination of the Torrance-Sparrow and Oren-Nayar facet models. There is no theoretical justification for such ad hoc combinations, and so they are probably best viewed as empirical fitting functions which work because they contain a sufficient number of free parameters. Newell and Keski-Kuha (1996) have looked at extreme-ultraviolet (UV) scatterers and baffles of interest to astronomers. Perhaps not surprisingly, materials which perform well in the IR and visible spectral ranges often make poor choices for the UV. Similarly, Watkins et al. (1993) made $s$- and $p$-polarized measurements of high-reflectance dielectric coatings using laser sources and an InGaAs or photomultiplier detector in a type (ii) configuration. Superpolishing of the substrates can reduce the BRDF, but generally only at the design wavelength—scatter characteristics at different wavelengths tend to be uncorrelated. Bickel et al. (1987; also see Iafelice and Bickel, 1987) have measured the complete Mueller matrices for smooth and distorted metal surfaces by periodically modulating the incident polarization state from a laser and measuring the fundamental and second harmonic of the scattered light in a type (ii) setup; this demonstrates the sensitivity of polarimetric optical scattering to surface perturbations. An alternative technique for finding the Mueller matrix uses a fixed linear polarizer and a variable retarder in both the source and receiver arms (Sornsin and Chipman, 1996); an integrating sphere scrambles the analyzed state to compensate for any polarization sensitivity of the detector. Burnell et al. (1994) have studied oxidized nickel collected from the interior of furnace tubes for the purposes of distinguishing emitted and reflected radiation in pyrometric measurements. After finding that the data could not be fit with physically realistic values of the parameters in the Torrance-Sparrow model, they also modified it in an ad hoc fashion to give a function which worked better but could not be interpreted physically. This is reminiscent of the approach used to devise the coupled model (Shirley et al., 1997) or the (similar) modified Beard-Maxwell model (Ellis, 1996a) for glossy surfaces.
References
H.C. Van de Hulst (1957), Light Scattering by Small Particles (John Wiley, NY).