

Advanced Real-Time Rendering: Beyond Reflectance

Naty Hoffman
Naughty Dog, Inc.
naty@naughtydog.com

Peter-Pike Sloan
Microsoft
ppsloan@windows.microsoft.com

Dan Baker
Microsoft
danbaker@windows.microsoft.com

Introduction

Reality is the benchmark against which computer-generated imagery has always been measured. With few exceptions, rendering processes used in games, films and academic research attempt to create images which resemble photographs, even when the depicted objects themselves are purposely stylized.

Research in computer graphics has historically emphasized physically-based rendering: the simulation of underlying physical phenomena in an attempt to create images which are identical to real ones. In animated films the emphasis is much more on enabling artistic control, eliminating visual artifacts and reducing rendering time, than on simulating physical processes. Rather than exactly matching real images, the target is to *resemble* them; verisimilitude, rather than reproduction, is the goal. In other words, the desired target is not photorealism but rather the *illusion* of photorealism. In games, the target is the same (although harder to reach), and the emphasis on physical correctness is even lower. As available power increases, game developers aspire to create real-time renderings on par with the visual quality found in feature films. It is natural to look to rendering techniques used in film production and to continue to place a low value on physical correctness. However, there are important differences between film and game rendering that deserve closer inspection.

In animated films, each shot is painstakingly crafted to look “just right”. The scene is rendered in dozens – sometimes hundreds – of separate layers, with independent lighting settings for each layer. These layers are then composited together, and extensive post-processing is applied. To a large extent, the computer is a tool which is used much as an artist uses a paintbrush – to “paint” each shot in a certain way. Under these conditions, the illusion of photorealism can be achieved by tweaking each shot until it *appears* real, even if the rendering techniques used bear no relation to physical reality.

In game rendering, the lighting, camera and scene can all change unpredictably in response to actions taken by the player. Obviously, no manual compositing, processing, or tweaking can take place. The final rendered image is not directly controlled by the game developer; the developers’ visual design is applied indirectly by controlling the players’ environment and the techniques used to render that environment. As such, it is much harder to achieve verisimilitude with rendering techniques which diverge in *significant* ways from physical reality. Sooner or later there will be a camera angle, lighting condition, or some other effect which will break the illusion of photorealism. Using reality-based rendering techniques also helps achieve consistency as lighting and camera settings change from frame to frame.

We are not, however, advocating absolute physical correctness in game rendering; such a thing does not exist in any case. Note the use of the word “significant” in the previous paragraph. It is possible to diverge from physical reality in many ways which are *not* significant to the goal of achieving an illusion of photorealism. No matter how powerful graphics hardware becomes, we will always have to take many shortcuts to render complex scenes at interactive rates; a basic understanding of the physical principles of rendering and human perception are very useful when deciding which shortcuts to take. Even when purposely stylized or non-photorealistic effects are desired, knowing the principles involved helps to diverge from reality in a controlled and purposeful manner; it has often been said that it is easier to break the rules if you know them well first.

This paper accompanies our lecture “Advanced Real-Time Rendering: Beyond Reflectance”. It focuses on a qualitative description of physical phenomena and creates the context for the discussion of various rendering techniques (which will not be covered here).

The Scene

The matter in a rendered scene can be divided into three groups. Some matter is actively emitting light. Secondly, some matter is interacting passively with the emitted light. Finally, there is a special bit of matter (which also interacts passively with the light): the sensor we are using to capture the scene. This may be a camera or the human visual system. The interaction of light with the sensor produces the final image of the scene. We want to understand how the light is emitted and how it interacts with the scene and the sensor.

The Nature of Light

Depending on circumstances, light can appear to behave as a wave or as a particle. Most phenomena relevant to rendering can be discussed using either a wave or particle model; we shall use whichever one is most appropriate to the phenomenon under discussion.

Light is electromagnetic radiant energy, but not all such energy impinging on the sensor will be registered as part of the image – only the subset to which the human visual system is sensitive. Figure 1 shows the human visual sensitivity curve in relation to the entire electromagnetic spectrum. Note that, due to the dual nature of light, there are two numbers used to quantify the electromagnetic radiation. The numbers at the top of the scale represent the energy of a single photon. This value is measured in electron-Volts (eV) which is defined as the amount of energy required to move a single electron through a potential difference of one Volt. The numbers along the bottom of the scale represent the wavelength of the light. This value is measured in meters. Note that wavelength and energy are inversely related.

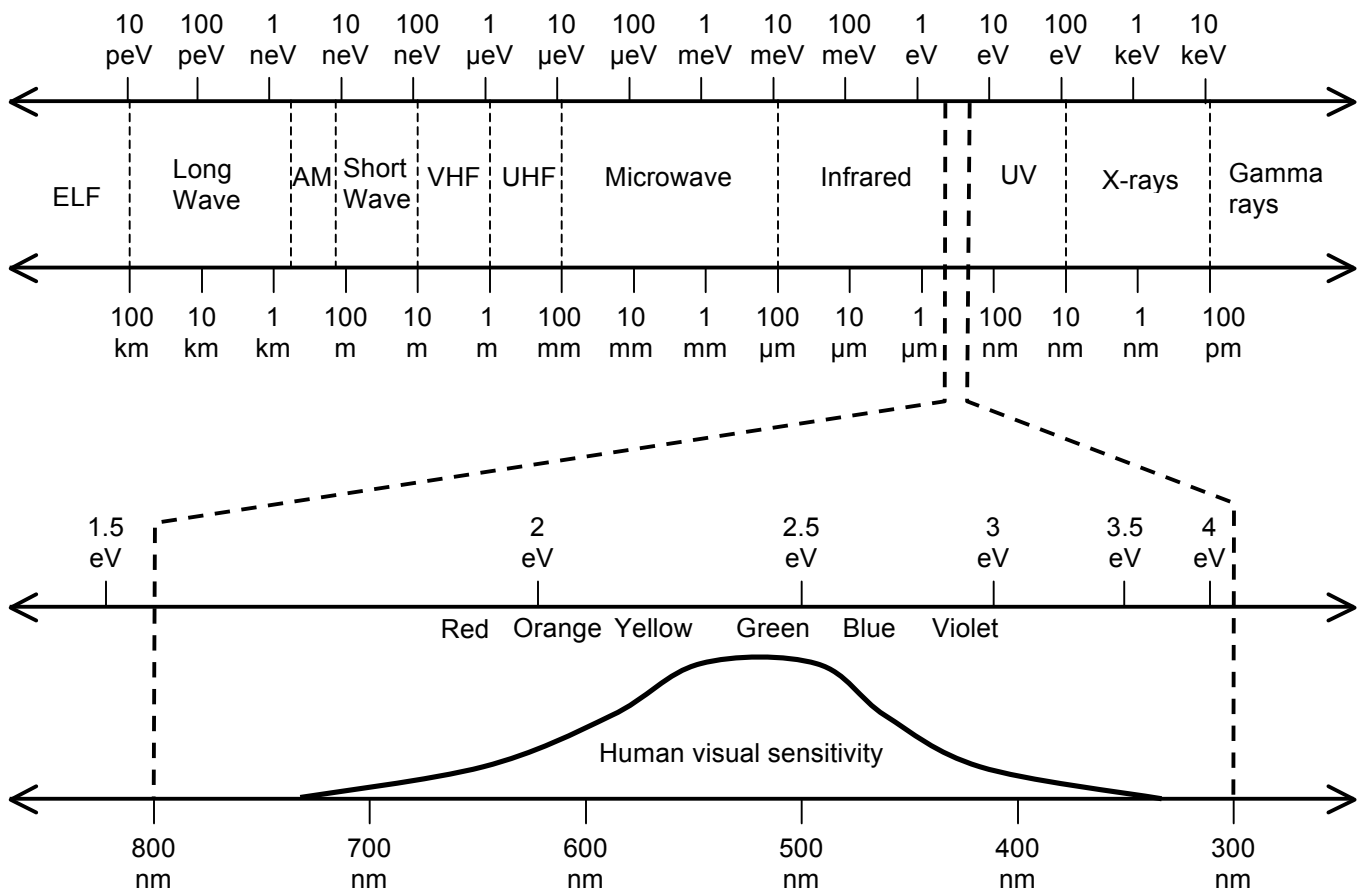


Figure 1: Electromagnetic spectrum and human visual sensitivity

The Particle Model of Light

Light particles are called *photons*. Each photon has momentum and energy proportional to its frequency, but has no mass or electric charge. In a vacuum a photon travels at velocity c (approximately 300,000 km/sec), which is faster than any particle with mass.

Photons also spin about the axis defined by their direction of travel. This spin can be right- or left-handed, and it gives them angular momentum.

The Wave Model of Light

The wave model of light is considerably more detailed. Light is a wave propagating through two coupled fields: electric and magnetic. These *electromagnetic waves* are described by *Maxwell's equations* [Maxwell1865]. Light propagates as a *transverse wave*; the constituent field vectors are orthogonal to the direction of propagation and to each other. The part of the wave which is at a given phase travels at the *phase velocity*, which is c in a vacuum. In Figure 2, red waves represent the electric field and blue waves represent the magnetic field:

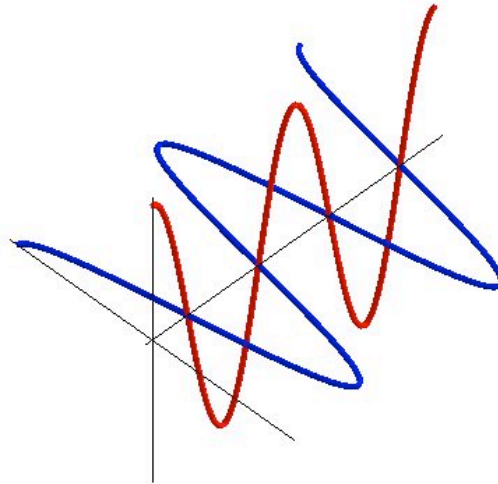


Figure 2: Electromagnetic transverse wave

Note that although this plot appears to be a 3D spatial plot it actually isn't. The longest of the three axes (along which the wave is propagating) is spatial. However, the other two axes do not indicate the spatial extent of the wave, but rather the electrical and magnetic field strength in the other two spatial dimensions. This is also true of similar plots elsewhere in this document.

The light wave in Figure 2 is a *simple harmonic wave* with a single wavelength. Such light is *monochromatic* and *coherent*. Coherent light is emitted by lasers, but it is not the most common type of light. Light is typically composed of a superposition of waves with many different wavelengths. That is, it is *polychromatic* and *incoherent*.

Another important property of light is its *polarization state*. To illustrate, we will look at a simple harmonic plane as shown in Figure 2, and examine the behavior over time of the electric field vector at a given point. (The magnetic field vector always exists at a fixed relationship to the electric field vector, so nothing is gained from showing it again). We will see that the electric field vector moves in the *plane of vibration*, which is the plane perpendicular to the direction of wave propagation. There are three possible cases for this motion, as shown in Figures 3-5):

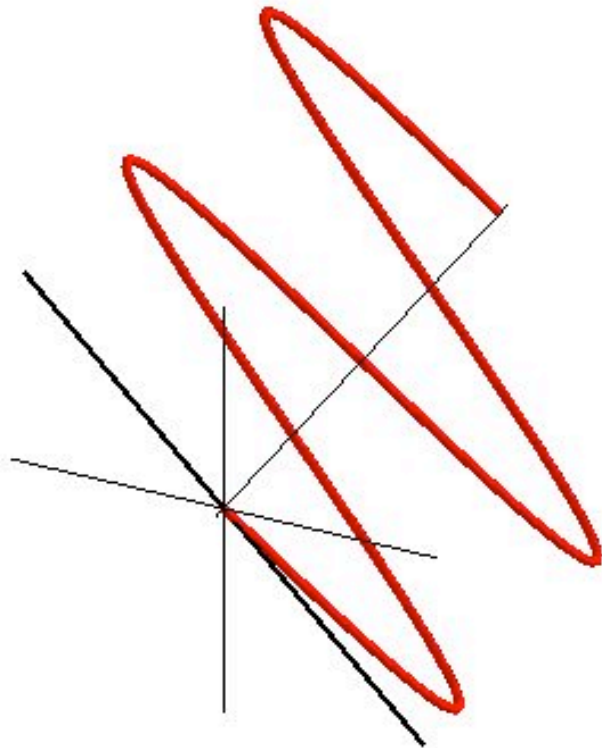


Figure 3: Linearly polarized wave

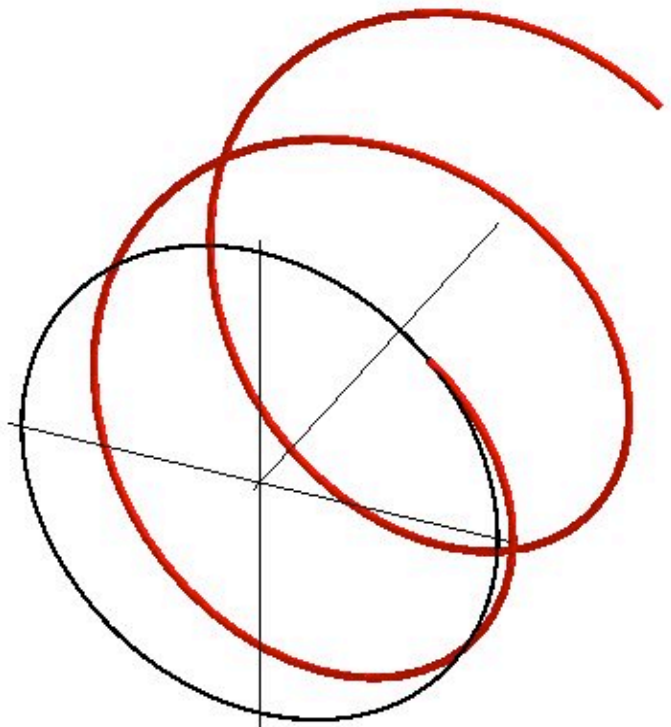


Figure 4: Circularly polarized wave

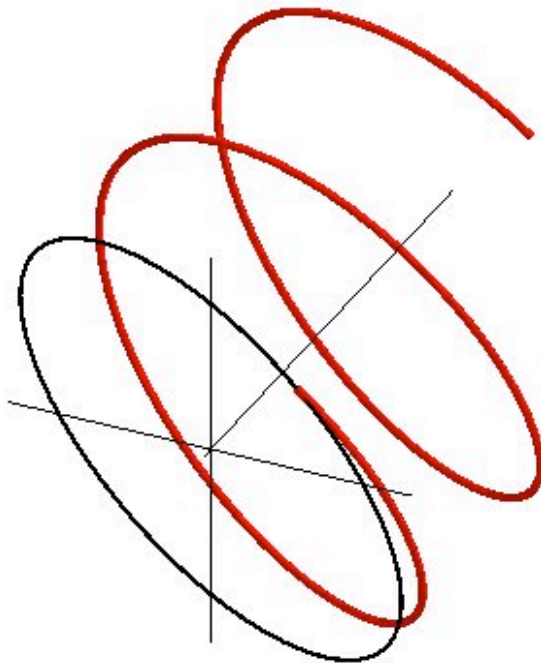


Figure 5: Elliptically polarized wave

In these figures, the wave travels to the lower left and out of the plane of the picture along the axis of propagation; the other two axes define the plane of vibration. The thin black line shows

the path traced over time by the tip of the electric field vector at the origin. The thick red line shows the path traced by the tip of the electric field vector in space, which moves towards us at velocity c .

In Figure 3, the electric field vector at the origin moves back and forth along a line; the wave as a whole is contained within the *plane of polarization*. This light is *linearly polarized*. In Figure 4, we see that the tip of the electric field vector traces a circle. This light is *circularly polarized*. In this figure the electric field vector sweeps in the counter-clockwise direction. This is *left-handed circular polarization*. If the field vector rotated clockwise, the light would be *right-handed circularly polarized*. In Figure 5, we see that the electric field vector rotates counter-clockwise as in Figure 4, but it traces an ellipse rather than a circle. This light is *elliptically polarized*. This is the most general polarization state for coherent light, and it includes the other two as special cases. For this reason we will refer to the path traced by the tip of the electric field vector as the *polarization ellipse*, even for light that is linearly or circularly polarized. Note that circularly and elliptically polarized light has no plane of polarization. It is also important to note that any coherent wave can be decomposed into two linearly polarized components. Similarly, any coherent wave can be decomposed into two circularly polarized components, one left-handed and one right-handed.

Coherent light always exists in one of these three polarization states. However, incoherent light can be *polarized*, *unpolarized* or *partially polarized*. If it is polarized, it can exist in any of the same polarization states as coherent light. (The tip of the field vector may not sweep as smoothly through the cross-section curve, however). If the light is unpolarized, the field vectors may point in any direction. Partially polarized incoherent light (the most general case) can always be decomposed into two components: one fully polarized and one fully unpolarized. The polarization state of incoherent light can thus be fully defined by the polarization ellipse and the ratio of polarized intensity to total light intensity.

The Quantum Model of Light

Quantum field theory combines the classical wave and particle views of light. In this theory, photons are considered to be electromagnetic “wave packets” which obey Maxwell’s equations. These packets are formed by the superposition of waves in a continuous range of frequencies such that the resulting wave has a limited spatial extent. Figure 6 shows a simplified two-dimensional drawing of a wave packet.

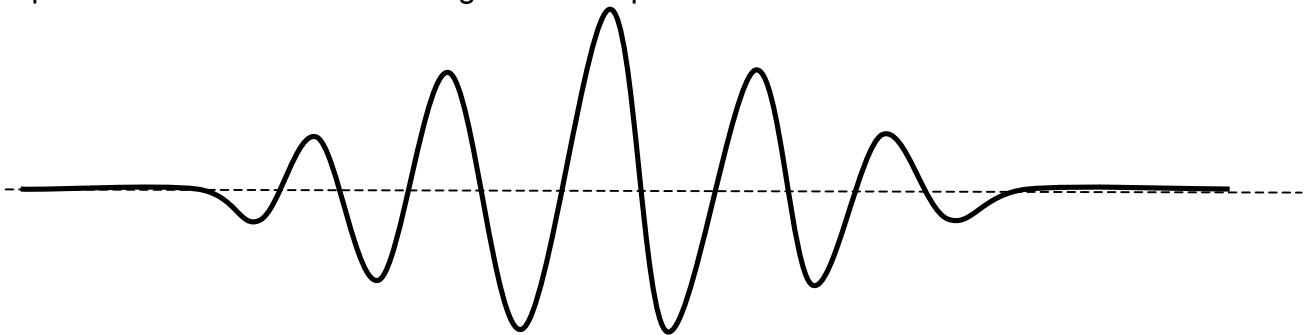


Figure 6: Electromagnetic wave packet

The frequency content of this wave packet and its spatial extent are inversely related. If we reduce the range of frequencies contained in the wave packet it will spread out in space until (in the limit) it becomes a monochromatic harmonic wave of infinite extent. To decrease the spatial extent of the wave packet, the range of frequencies it contains must be expanded until in the limit it is an infinitesimally wide *delta function* containing all frequencies. If we look at this wave packet as a photon, remembering that the momentum of a photon is proportional to its frequency, we see that it has indeterminate position and momentum, and these indeterminacies are inversely related. This is the *Heisenberg uncertainty principle* [Heisenberg1927] as applied to photons. The spin of a photon can take on one of two values: one is the negation of the other. These two possibilities are related to the right-handed and left-handed circular polarization of the light. This is a very rough description of quantum field theory; there are many subtleties that which we do not have space to cover here.

The Interaction of Light with Matter

The interaction of light with matter is most relevant to rendering. We will discuss three cases of light interaction: light emission, light propagation through a medium, and light interaction with the boundary between two media. The components of matter which most commonly interact with light are electrons. Note that the ratio of magnetic force to electric force on an electron is the ratio between its velocity and c ; the interaction of light with matter is mostly electric in nature.

Light Emission

When atoms are *excited* to an energy level above their *rest state* (lowest possible energy level), they may drop back to a lower energy level, releasing a photon in the process. More energetic atoms release higher-energy, and thus lower-wavelength, photons. This is why incandescent objects glow red-yellow to blue-white as their heat increases. Different light sources emit light with different spectral distributions. For example, the sun emits light with a wide spectral distribution, but most of its energy is concentrated in the visible portion of the spectrum.

Light Propagation through a Medium

Light propagates through a vacuum with a phase velocity of c . It propagates through other media at a different, usually lower, phase velocity. The amount by which light is slowed is called the *index of refraction*. For example a medium with an index of refraction of 2 will cause light to propagate at $0.5c$. The index of refraction is 1 in a vacuum by definition. In some cases, a medium can have an index less than 1, which will cause electromagnetic waves to have a phase velocity greater than c . This does not violate relativity since information does not necessarily travel at the phase velocity. This change in phase velocity affects the lights' wavelength but not its frequency; this shows that frequency is a more fundamental property than wavelength. In addition to slowing the phase velocity, some media will also cause the lights' amplitude to be attenuated; this is called *absorption*. Maxwell's equations indicate that in a homogeneous medium the attenuation is exponential with distance. The fractional attenuation per unit distance is the *extinction coefficient* of the medium. This value is often

combined with the index of refraction into a single *complex index of refraction*, where the imaginary part is proportional to the extinction coefficient.

The refractive index of a medium is not a constant; it is a function of frequency. This spectral dependence is called *dispersion*. In transparent media the refractive index usually increases with frequency (*normal dispersion*). When the refractive index decreases with frequency, it is called *anomalous dispersion*.

There are two main types of media: those which conduct electricity and those which do not (dielectrics). Conductive media are usually highly absorbent in the visual band. This is because they have a "sea" of free electrons that can change energy levels freely. Dielectrics are usually transparent (non-absorbent) at most frequencies. Most of their electrons are tightly bound to atoms, which limits the photons they can absorb. Bound electrons can only absorb photons which have the right energy to move their bound atoms between two allowed energy states (the quantum view) or cause them to vibrate at their resonant frequency (the classical view). For most dielectrics, these *absorption bands* are outside the visual spectrum. One notable exception is water, which absorbs some light in the red part of the spectrum. (This gives water its characteristic blue color.) Most colored gemstones are examples of *selective absorption* as well. Semiconductors have absorption properties that lie between those of dielectrics and conductors.

When a photon is absorbed its energy must go somewhere. In the case of a conducting medium, the absorbing electrons will usually induce alternating electric currents which will immediately cause an identical photon to be re-emitted. In the case of a dielectric medium, the absorbed photon will excite an atom or molecule into a higher energy state. In some cases, the excited atom may step down to a lower energy level and re-emit a photon, either immediately (*fluorescence*) or after a time delay (*phosphorescence*). This photon may not be of the same frequency as the one that was originally absorbed.

If the medium is *optically isotropic*, then its properties are the same regardless of the direction of light propagation. Some media are *optically anisotropic*; they have some directional structure which causes them to have different indices of refraction and/or extinction coefficients depending on the direction of propagation. This also causes the medium to affect light differently depending on its linear polarization. If the index of refraction varies based on direction of propagation or linear polarization, then the medium is *birefringent* or *linearly birefringent*. If the extinction coefficient varies based on direction of propagation or linear polarization, then the medium is *dichroic* or *linearly dichroic*. Many crystals and minerals are optically anisotropic.

A medium is known as *optically active* if it actively changes the polarization state of light passing through it. Birefringent and dichroic media also change polarization state, but they do so in a passive manner by selectively absorbing or retarding light of certain linear polarizations. Optically active media are sensitive to circular polarization and can rotate the plane of polarization for linearly polarized light. They can also change linearly polarized light into circularly polarized light and vice versa. Optically active media which have a different index of refraction for right- and left-handed circularly polarized light are called *circularly birefringent*.

Media which have a different extinction coefficient for right- and left-handed circularly polarized light are called *circularly dichroic*. Most optically active media are organic and composed of molecules with complex, twisting shapes.

Light Interaction with a Boundary Between Two Media

When light hits a smooth boundary between two media, part of it is returned back into the originating medium (*reflection*) and the rest of it continues on into the new medium after changing direction (*refraction*). Both effects are a consequence of Maxwell's equations in the case of a refractive index discontinuity. The resulting directions of the reflected and refracted light have been known since the 17th century (see Figure 7):

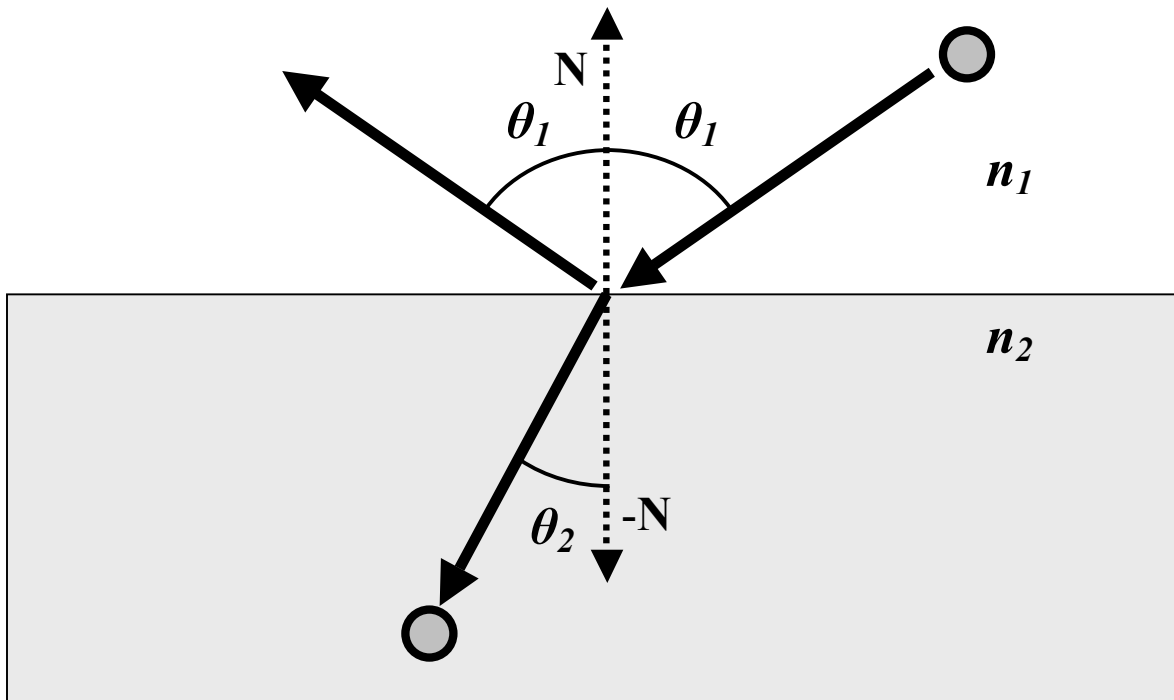


Figure 7: Directions of reflection and refraction

The reflected light vector and the incident light vector form equal angles with the surface normal, and all three vectors are coplanar. The refracted light vector is also coplanar with the incident light vector and the surface normal, but its angle with the normal depends on the (real) indices of refraction of both media and is described by *Snell's law*. (The form currently in use was actually published by Descartes in [Descartes1637], but Snell discovered it several years previously.) Snell's law states that the ratio between the sines of the angles θ_1 and θ_2 is the inverse of the ratio between the corresponding real indices of refraction n_1 and n_2 . This bending of the light also has a remarkable property which is related to how the index of refraction affects the phase velocity of the light. If we look at any two points, one on either side of the boundary between media (see the circles in Figure 7), then the bent path caused by refraction is the quickest possible path between these two points given the velocity of light in the respective media. This is an example of *Fermat's principle* which states that light always travels along the quickest path between two points (more precisely, the travel time has to be a

local minimum or maximum). Qualitatively, Snell's law means that when light travels from a medium with a lower index of refraction to one with a higher index of refraction, its direction is bent towards the line parallel to the surface normal. As the ratio between the two indices of refraction increases the bending effect also increases. When light travels from a medium with a higher index of refraction to one with a lower index of refraction, it is bent *away* from this axis. At a certain angle, the light will be parallel to the surface; this is called the *critical angle*. Light which strikes the media boundary at angles outside the critical angle undergoes *total internal reflection*. That is, there is no refraction; all light is reflected back into the denser medium. In Figure 8 the solid arrows represent an incident light beam inside the critical angle and its resulting reflected and refracted beams; the dashed arrows represent an incident light beam outside the critical angle which undergoes total internal reflection. The critical angle is marked by a dark circular wedge.

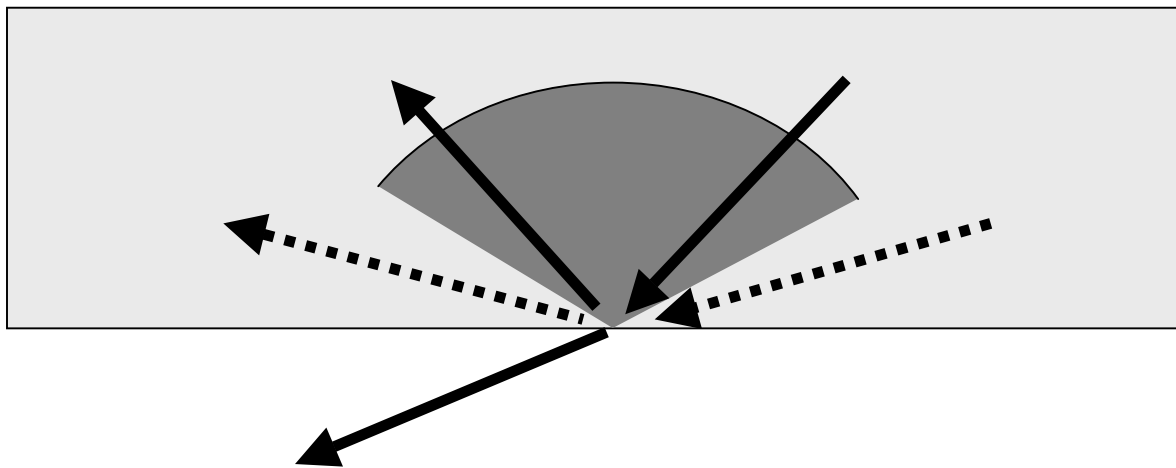


Figure 8: Critical angle and total internal reflection

As mentioned in the previous section, the index of refraction of most media is spectrally dependent. This causes the angle of refraction to vary depending on the lights' wavelength. A beam of polychromatic light will disperse into its constituent wavelengths upon passing into a dispersive medium at an oblique angle (see Figure 9).

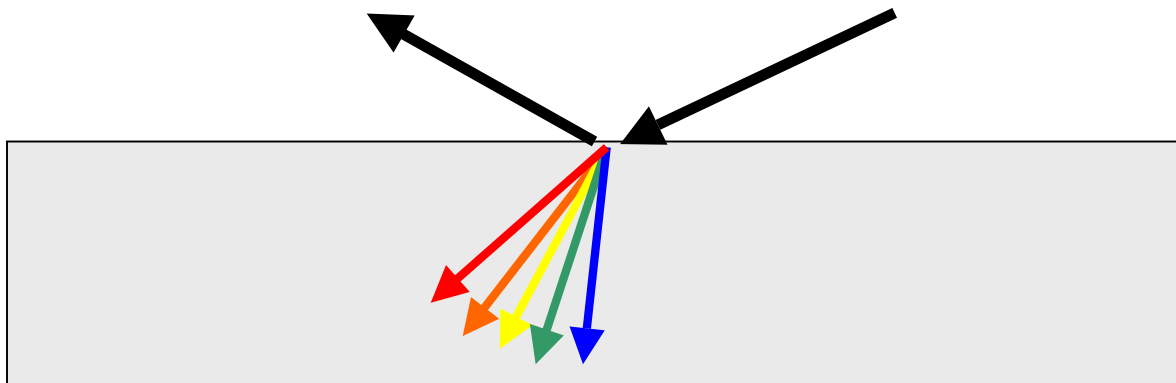


Figure 9: Dispersion

While the direction of the reflected and refracted light is determined by the law of reflection and Snell's law, the lights' intensity and polarization is described by the *Fresnel equations* [Fresnel1822]. These equations are a special case of Maxwell's equations for the boundary between two media. The full equations are somewhat complex, so we will only summarize their most relevant results. In the case of light traveling from a vacuum or air (in which the refractive index is 1 or almost 1) to some denser media, we can infer the following points from the Fresnel equations and common material properties:

- As the angle of incidence tends to 90° , the reflectance tends to 1 for all frequencies, polarizations and materials. That is, all light is reflected at very glancing angles.
- At normal incidence (angle of incidence is 0° - the incident light is perpendicular to the surface), the reflectance depends on the complex index of refraction (a spectral quantity). For dielectrics the normal-incidence reflectance tends to be low and constant throughout the visible spectrum; most range between 2% (water) and 17% (diamond). For metals it tends to be higher (above 50% and even 75%) and often varies throughout the visible spectrum.
- At intermediate angles, the reflectance depends on whether the light is polarized parallel to the surface or perpendicular to it. The parallel-polarized reflectance increases monotonically from the normal-incidence reflectance to 1 as the angle of incidence goes from 0° to 90° . The perpendicular-polarized reflectance first dips below the normal-incidence reflectance, then reaches a minimum value, and finally tends back up to 1 as the angle of incidence goes from 0° to 90° . This minimum value is 0 for dielectrics, so dielectrics tend to polarize light reflected from them quite strongly. For metals, the minimum value is not very low, so the resulting polarization is not that strong. The effects of the "dip" are commonly seen in metals as a slight color shift. For example, copper has reddish reflectance at normal incidence, a bluish color at intermediate angles, and a white color at glancing angles.

From this we see that most dielectrics are quite transparent to light, especially at normal incidence. Much of the incident light penetrates into the medium where it can be affected by phenomena such as selective absorption. The surface reflections from dielectrics are usually colorless. Metals, on the other hand, are highly reflective and their surface reflections are often colored. Metals quickly absorb any refracted light. Finally, dielectrics tend to polarize reflected light, which emphasizes the components parallel to their surface.

In some cases, boundary effects predominate over interior effects. For example, if a substance is ground into a fine powder, there is a lot of surface area but little interior volume with which light can interact. For this reason, substances which have color resulting only from interior effects (such as gemstones) will be colorless in the form of a fine powder. A similar effect can be observed on the foam that forms on liquids. Even if a liquid is strongly colored, its foam is usually colorless. This does not happen with materials which have color due to surface effects. For example, copper which is ground into a fine powder retains its reddish color.

When the boundary is not smooth but has structure of a size comparable to the light wavelength various other effects occur, mostly related to constructive and destructive *interference* among light waves. The shifting colors displayed by certain types of bird feathers

(most notably peacock tails), beetles, soap bubbles, oil puddles, etc. are examples of this phenomenon.

Conclusions

The physics of light and its interaction with matter are intimately connected to rendering. Understanding the underlying physical principles can help us achieve better quality rendering. Even when approximations are used which ignore certain physical phenomena, it is helpful to understand the phenomena involved so that educated choices of approximation can be made.

References

[Descartes1637] DESCARTES, R. *La Dioptrique* (Leiden, 1637).

[Fresnel1822] FRESNEL, A. J. Extrait d'un m'emoire sur la double r'efraction particul`ere que pr'esente le cristal de roche dans la direction de son axe. Bulletin des Sciences par la Societe Philomatique de Paris (1822), pp. 191-198.

[Heisenberg1927] HEISENBERG, W. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Zeitschrift für Physik 43 (1927), pp. 172-198.

[Maxwell1865] MAXWELL, J. C. A Dynamical Theory of the Electromagnetic Field. Philosophical Transactions of the Royal Society of London 155 (1865), pp. 459-512.