Sphere Intersection

We know that a sphere of radius r is the set of all points a distance r from sphere's center. In vector notation, this is

$$(\mathbf{P} - \mathbf{O}_S) \cdot (\mathbf{P} - \mathbf{O}_S) = r^2$$

where $\mathbf{P}$ is a vector touching a point on the sphere and $\mathbf{O}_S$ is the vector describing the sphere's center.

We describe a ray via

$$\mathbf{P}(t) = \mathbf{O}_R + \mathbf{D}t$$

where $t \in \mathbb{R}$. Substituting this into the implicit equation for our sphere yields

$$\begin{aligned} r^2 &= (\mathbf{P}(t) - \mathbf{O}_S) \cdot (\mathbf{P}(t) - \mathbf{O}_S), \\ &= (\mathbf{O}_R + \mathbf{D}t - \mathbf{O}_S) \cdot (\mathbf{O}_R + \mathbf{D}t - \mathbf{O}_S), \\ &= \big(\mathbf{D}t + (\mathbf{O}_R - \mathbf{O}_S)\big) \cdot \big(\mathbf{D}t + (\mathbf{O}_R - \mathbf{O}_S)\big), \\ &= (\mathbf{D} \cdot \mathbf{D})t^2 + 2\mathbf{D} \cdot (\mathbf{O}_R - \mathbf{O}_S)t + (\mathbf{O}_R - \mathbf{O}_S) \cdot (\mathbf{O}_R - \mathbf{O}_S). \end{aligned}$$

We can define $\mathbf{U} = \mathbf{O}_R - \mathbf{O}_S$, collect everything on one side, and group terms by orders of t to arrive at

$$(\mathbf{D} \cdot \mathbf{D})t^2 + (2\mathbf{D} \cdot \mathbf{U})t + \big(\mathbf{U} \cdot \mathbf{U} - r^2\big) = 0$$

which is a simple quadratic equation in t. Its solution is

$$t = \frac{-2\mathbf{D} \cdot \mathbf{U} \pm \sqrt{\big(2\mathbf{D} \cdot \mathbf{U}\big)^2 - 4(\mathbf{D} \cdot \mathbf{D})(\mathbf{U} \cdot \mathbf{U} - r^2)}}{2\mathbf{D} \cdot \mathbf{D}}.$$

or, removing a common factor of 2,

$$t = \frac{-\mathbf{D} \cdot \mathbf{U} \pm \sqrt{\big(\mathbf{D} \cdot \mathbf{U}\big)^2 - (\mathbf{D} \cdot \mathbf{D})(\mathbf{U} \cdot \mathbf{U} - r^2)}}{\mathbf{D} \cdot \mathbf{D}}.$$

Evidently, the ray intersects our sphere when the discriminant of this equation is greater than or equal to 0, i.e.

$$\big(\mathbf{D} \cdot \mathbf{U}\big)^2 - (\mathbf{D} \cdot \mathbf{D})(\mathbf{U} \cdot \mathbf{U} - r^2) \geq 0$$

If this condition is met, the value of t corresponding to the first intersection is the smallest value that is still greater than 0.

Importance Sampling

For a function $f(x)$ of a random variable X, its expected value is given by

$$\mathbb{E}_p[f(X)] = \int_\Omega f(x) p_X(x) \, \mathrm{d}x$$

where $p_X(x)$ is the probability density function (PDF) of X. A statistical way to evaluate the expected value is given by the sample mean, defined by

$$\mathbb{E}_p[f(X)] = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$

where each x_i is a realization of X drawn from the distribution $p_X(x)$. We often want to evaluate integrals of the form $\int_\Omega f(x) \, \mathrm{d}x$, which we can do by taking the expected value of $f(x)/p_X(x)$ instead of just $f(x)$. Truncating the sample mean at a finite N, we see the correspondence

$$\int_\Omega f(x) \, \mathrm{d}x \approx \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p_X(x_i)}$$

If the samples we use are not a good fit to f, we run the risk of adding up a lot of terms that contribute very little to the final result. We can work around this by including a simple multiplicative factor of 1 in the expected value in the form of $q_X(x)/q_X(x)$, where $q_X(x)$ is a PDF that better matches f. Specifically, we see

$$\mathbb{E}_p\Big[\frac{f(X)}{p_X(X)}\frac{q_X(X)}{q_X(X)}\Big] = \int_\Omega \frac{f(x)}{p_X(x)}\frac{q_X(x)}{q_X(x)}p_X(x)\,\mathrm{d}x = \int_\Omega \frac{f(x)}{q_X(x)}q_X(x)\,\mathrm{d}x = \mathbb{E}_q\Big[\frac{f(X)}{q_X(X)}\Big]$$

which gives us the correspondence

$$\int_\Omega f(x)\,\mathrm{d}x \approx \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{q_X(x_i)}$$

where, in this case, each x_i is sampled from the distribution defined by $q_X(x)$. There are, of course, restrictions on the possible choices of $q_X(x)$, but they are surprisingly soft: we simply require that $q_X(x)$ be nonzero wherever $f(x)p_X(x)$ is nonzero.

In path tracing, we use this to estimate the indirect lighting present in the scene. Specifically, we take $f(x)$ to be the integrand of the rendering equation and N to be 1,

$$\int_\Omega f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|\,\mathrm{d}\hat{\omega}_i \approx \frac{f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|}{p_\omega(\hat{\omega}_i)}$$

Lambertian Reflectance

Recall the rendering equation,

$$L_o(\mathbf{x},\hat{\omega}_o,\lambda,t) = L_e(\mathbf{x},\hat{\omega}_o,\lambda,t) + \int_\Omega f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|\,\mathrm{d}\hat{\omega}_i$$

where $f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t)$ describes the bidirectional reflectance distribution function (BRDF) of the material under consideration and Ω is the hemisphere centered at the geometry's normal vector.

A diffuse, Lambertian material has an equal chance to scatter light in all directions. This implies that $f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t) = \rho_0$ where ρ_0 is a constant. Because of this, the direction of the incoming radiance does not matter and we may take $L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)$ outside of the integral. Assuming all light is reflected imposes the additional requirement that

$$\int_\Omega \rho_0|\hat{\omega}_i\cdot\mathbf{n}|\,\mathrm{d}\hat{\omega}_i = \rho_0\int_\Omega \cos\theta\,\mathrm{d}\hat{\omega}_i = 1$$

where θ is the angle between $\hat{\omega}_i$ and the normal. On the unit hemisphere, with ϕ describing the azimuthal angle and θ the altitude, this requirement becomes

$$\begin{aligned}
\rho_o\int_0^{2\pi}\int_0^{\pi/2}\cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi &= \frac{\rho_0}{2}\int_0^{2\pi}\int_0^{\pi/2}\sin(2\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi \\
&= \frac{\rho_0}{4}\int_0^{2\pi} -\cos(2\theta)\Big|_0^{\pi/2}\,\mathrm{d}\phi \\
&= \frac{\rho_0}{2}\int_0^{2\pi}\mathrm{d}\phi \\
&= \rho_0\pi \\
&= 1
\end{aligned}$$

i.e. max reflectance occurs when $\rho_0 = 1/\pi$. If we allow this to vary by wavelength via $f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t) = \rho(\lambda)/\pi$ with $\rho(\lambda)\in[0,1]$, the indirect part of the path tracing equation becomes

$$\frac{\rho(\lambda)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|}{\pi p_\omega(\hat{\omega}_i)}$$

If we sample directions uniformly in the unit hemisphere, for example, $p_\omega(\hat{\omega}_i) = 1/2\pi$ and the above simplifies to

$$2\rho(\lambda)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|$$

Uniformly Sampling the Unit Sphere (and Unit Hemisphere)

The easiest way to generate uniformly spaced points on the unit sphere is to build a cumulative distribution function (CDF), then use inverse transform sampling. To put it another way, suppose we want to generate samples from a PDF $p_X(x) = \mathrm{d}F_X(x)/\mathrm{d}x$ but only have access to U, a random variable uniformly distributed between 0 and 1. An obvious way to map from this domain to the range of X is to define $Y = F_X^{-1}(U)$. The CDF of this new random variable is

$$F_Y(y) = \Pr(Y \leq y) = \Pr(F_X^{-1}(U) \leq y) = \Pr(U \leq F_X(y)) = F_X(y),$$

where we have used the fact that all CDFs are monotonic and that $\Pr(U \leq u) = u$ for a random variable uniformly distributed between 0 and 1. Given the uniqueness of the derivative, we can be confident that Y has the same distribution as X.

Moving on to the problem at hand, the PDF of the unit sphere with regard to solid angle Ω must be

$$p_\Omega(\Omega) = \frac{1}{4\pi}$$

We can, of course, write this in ϕ-θ space as

$$p_\Omega(\Omega)\,\mathrm{d}\Omega = p_{\theta,\phi}(\theta,\phi)\,\mathrm{d}\theta\mathrm{d}\phi$$

where, upon identifying $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$, we find

$$p_{\theta,\phi}(\theta,\phi) = \frac{\sin\theta}{4\pi}$$

We can integrate over each variable to obtain the PDF of the other,

$$p_\theta(\theta) = \int_0^{2\pi} \frac{\sin\theta}{4\pi}\mathrm{d}\phi = \frac{\sin\theta}{2}$$
$$p_\phi(\phi) = \int_0^{\pi} \frac{\sin\theta}{4\pi}\mathrm{d}\theta = \frac{1}{2\pi}$$

which can then be used to find their respective CDFs

$$F_\theta(\theta) = \int_0^{\theta} \frac{\sin\theta'}{2}\mathrm{d}\theta' = \frac{1-\cos\theta}{2}$$
$$F_\phi(\phi) = \int_0^{\phi} \frac{\mathrm{d}\phi'}{2\pi} = \frac{\phi}{2\pi}$$

whose inverse functions are

$$F_\theta^{-1}(u) = \cos^{-1}(1-2u) \equiv \theta$$
$$F_\phi^{-1}(v) = 2\pi v \equiv \phi$$

We can substitute these into the conversions from Cartesian to spherical coordinates to get

$$x = \sin\theta\cos\phi = \sqrt{1-z^2}\cos(2\pi v)$$
$$y = \sin\theta\sin\phi = \sqrt{1-z^2}\sin(2\pi v)$$
$$z = \cos\theta = 1-2u$$

Another method for uniformly sampling the unit sphere is to draw random values from the unit normal distribution for each component of a vector. One then simply needs to normalize the resulting vector. This works because the combined Gaussian PDF is rotationally invariant, and thus must correspond to a uniform distribution on the sphere.

Yet another method is the so-called 'rejection' method. In this, we draw values for the vector's components from a random variable uniformly distributed between -1 and 1. If the vector lies within the unit ball, we normalize it. Otherwise, we repeat the process.

After we have setup a method to sample the unit sphere, sampling the unit hemisphere is just as easy. We simply dot the sampled vector with the desired normal vector: if this result is less than 0, we negate the sampled vector.

Uniformly Sampling the Unit Disc

We follow a similar strategy to above. Over the unit disc, a uniform sampling strategy must yield

$$p_A(A) = \frac{1}{\pi}.$$

Identifying $\mathrm{d}A = r\,\mathrm{d}r\,\mathrm{d}\phi$ and equating the above PDF with the equivalent PDF in r-ϕ spaces gives

$$p_A(A)\,\mathrm{d}A = \frac{r}{\pi}\mathrm{d}r\,\mathrm{d}\phi = p_{r,\phi}(r,\phi)\,\mathrm{d}r\,\mathrm{d}\phi,$$

i.e. $p_{r,\phi}(r,\phi) = r/\pi$. The PDF of each variable is

$$p_r(r) = \int_0^{2\pi} \frac{r}{\pi}\mathrm{d}\phi = 2r$$
$$p_\phi(\phi) = \int_0^1 \frac{r}{\pi}\mathrm{d}r = \frac{1}{2\pi}$$

and so the corresponding CDFs are

$$F_r(r) = \int_0^r 2r'\,\mathrm{d}r' = r^2$$
$$F_\phi(\phi) = \int_0^\phi \frac{\mathrm{d}\phi'}{2\pi} = \frac{\phi}{2\pi}$$

and so the inverse CDFs are

$$F_r^{-1}(u) = \sqrt{u} \equiv r$$
$$F_\phi^{-1}(v) = 2\pi v \equiv \phi$$

In Cartesian coordinates, then, a randomly sampled point on the unit disc is given by

$$x = r\cos\phi = \sqrt{u}\cos(2\pi v)$$
$$y = r\sin\phi = \sqrt{u}\sin(2\pi v)$$

Cosine Weighted Sampling of the Unit Hemisphere

Looking back to the section on importance sampling, we see that sampling directions from a cosine weighted distribution has the advantage of removing the $|\hat{\omega}_i \cdot \mathbf{n}|$ term from the rendering equation. Let's figure out how to do this.

A cosine weighted distribution over the unit hemisphere should obey

$$p_\Omega(\Omega)\,\mathrm{d}\Omega = A\cos\theta\,\mathrm{d}\Omega = A\cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$$

where A is a normalization constant. We can find it by integrating over the unit hemisphere,

$$\begin{aligned}
\int_\Omega p_\Omega(\Omega)\,\mathrm{d}\Omega &= A\int_0^{2\pi}\int_0^{\pi/2} \cos\theta\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \\
&= \frac{A}{2}\int_0^{2\pi}\int_0^{\pi/2} \sin(2\theta)\,\mathrm{d}\theta\,\mathrm{d}\phi \\
&= -\frac{A}{2}\int_0^{2\pi} \frac{\cos(2\theta)}{2}\Big|_0^{\pi/2}\mathrm{d}\phi \\
&= \frac{A}{2}\int_0^{2\pi} \mathrm{d}\phi \\
&= A\pi \\
&= 1
\end{aligned}$$

i.e. $A = 1/\pi$. Using this, we may rewrite the PDF as

$$p_{\theta,\phi}(\theta,\phi) = \frac{\cos\theta\sin\theta}{\pi} = \frac{\sin(2\theta)}{2\pi}$$

The individual PDFs are

$$p_\theta(\theta) = \int_0^{2\pi} \frac{\sin(2\theta)}{2\pi}\mathrm{d}\phi = \sin(2\theta)$$
$$p_\phi(\phi) = \int_0^{\pi/2} \frac{\sin(2\theta)}{2\pi}\mathrm{d}\theta = -\frac{\cos(2\theta)}{4\pi}\Big|_0^{\pi/2} = \frac{1}{2\pi}$$

and their respective CDFs are

$$F_\theta(\theta) = \int_0^\theta \sin(2\theta')\,\mathrm{d}\theta' = \frac{1-\cos(2\theta)}{2}$$
$$F_\phi(\phi) = \int_0^\phi \frac{\mathrm{d}\phi'}{2\pi} = \frac{\phi}{2\pi}$$

The inverse CDFs are

$$F_\theta^{-1}(u) = \frac{\cos^{-1}(1-2u)}{2} \equiv \theta$$
$$F_\phi^{-1}(v) = 2\pi v \equiv \phi$$

Before expressing this in Cartesian coordinates, it is useful have the half-angle identities on hand

$$\cos\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1+\cos x}{2}} \qquad \sin\left(\frac{x}{2}\right) = \pm\sqrt{\frac{1-\cos x}{2}}$$

Using the above, we see

$$\cos\theta = \sqrt{1-u} \qquad \sin\theta = \sqrt{u}$$

where we have chosen the positive square root in each case so as to agree $\cos\theta$ and $\sin\theta$ when $\theta \in [0, \frac{\pi}{2}]$. Altogether this gives us

$$x = \sin\theta\cos\phi = \sqrt{u}\cos(2\pi v)$$
$$y = \sin\theta\sin\phi = \sqrt{u}\sin(2\pi v)$$
$$z = \cos\theta = \sqrt{1-u}$$

Comparing this to our strategy for randomly sampling the unit disc reveals that this is simply a projection of the former onto the unit hemisphere! Using this sampling method gives a particularly nice form to the integrand in the rendering equation in the case of a Lambertian surface,

$$\rho(\lambda)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)$$

Lack of Cosine Term in Specular Reflection

Specular reflection requires that the integral term in the rendering equation,

$$\int_\Omega f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|\,\mathrm{d}\hat{\omega}_i$$

conserve energy. In this particular case, the we know that $f_r(\mathbf{x},\hat{\omega}_i,\hat{\omega}_o,\lambda,t) \propto \delta(\mathbf{n}-(\hat{\omega}_i+\hat{\omega}_o)/2)$, and so, to conserve energy, we must have

$$\int_\Omega A\delta(\mathbf{n}-(\hat{\omega}_i+\hat{\omega}_o)/2)L_i(\mathbf{x},\hat{\omega}_i,\lambda,t)|\hat{\omega}_i\cdot\mathbf{n}|\,\mathrm{d}\hat{\omega}_i = L_i(\mathbf{x},2\mathbf{n}-\hat{\omega}_o,\lambda,t)$$

where A is a normalization constant. We see that the distribution must be

$$f_r(\mathbf{x}, \hat{\omega}_i, \hat{\omega}_o, \lambda, t) = \frac{\delta(\mathbf{n} - (\hat{\omega}_i + \hat{\omega_o})/2)}{|\hat{\omega}_i \cdot \mathbf{n}|}$$

and hence $|\hat{\omega}_i \cdot \mathbf{n}|$ does not appear in implementations of specular reflection (or indeed any delta function distribution).

Russian Roulette

Suppose we have an estimator F that we replace with an estimator F'

$$F' = \begin{cases} \frac{F}{p} & u \leq p \\ 0 & u > p \end{cases}$$

where u is a uniform random value between 0 and 1 and p is some valid probability. In words, F' returns F/p with chance p and 0 with chance $(1-p)$. This new estimator has the same expected value as F, because

$$\mathbb{E}[F'] = p \cdot \frac{\mathbb{E}[F]}{p} + (1-p) \cdot 0 = \mathbb{E}[F]$$

If we apply this to the indirect lighting integrand in our path tracer, we see that it allows the termination of a loop that would otherwise need to continue until hitting a light source—a process that could easily have no end. This has the advantage of allowing us to run our program in a finite time without incurring bias.

Reflection and Refraction

By simple geometry, we recognize a reflection about the surface normal must obey

$$\mathbf{n} = \frac{\hat{\omega}_o - \hat{\omega}_i}{2},$$

where $\hat{\omega}_o$ is the unit vector in the direction of the outward going ray and $\hat{\omega}_i$ is the unit vector in the direction of the inward coming ray. We can easily solve this for the outgoing ray,

$$\hat{\omega}_o = 2\mathbf{n} + \hat{\omega}_i.$$

Refraction is slightly more complicated. One may arrive at the famous Snell's law by minimizing the time it takes a light ray to traverse the distance between two points located in two homogenous mediums. It reads

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

where $n_i = c/v_i$ is the refractive index of each medium and θ_i is the angle the light ray makes with respect to the closest surface normal. When $\frac{n_1}{n_2} \sin\theta_1 > 1$, the above has no solutions and the ray must be reflected.

To find $\hat{\omega}_o$ in the case of refraction, it's easiest to split both direction vectors into components parallel and perpendicular to the surface normal. Since we are dealing with unit vectors, $\sin\theta_1 = \|\hat{\omega}_{i,\perp}\|$ and $\sin\theta_2 = \|\hat{\omega}_{o,\perp}\|$. With both direction vectors lying in the same plane, we can combine this with Snell's law to find

$$\hat{\omega}_{o,\perp} = \frac{n_1}{n_2}\hat{\omega}_{i,\perp} = \frac{n_1}{n_2}\big(\hat{\omega}_i - (\hat{\omega}_i \cdot \mathbf{n})\mathbf{n}\big).$$

We can get the parallel component of $\hat{\omega}_o$ through the Pythagorean theorem,

$$\hat{\omega}_{o,\|} = -\sqrt{1 - \|\hat{\omega}_{o,\perp}\|}\mathbf{n}$$

Both of these short derivations assume an outward facing surface normal. To deal with inward-pointing surface normals, we can either negate $\mathbf{n}$ in the intersection record or negate the instances of $\mathbf{n}$ in the above equations.

Fresnel Equations

When an electromagnetic wave traverses the boundary between two dielectric mediums, a portion of the light is reflected while the rest is transmitted. Ray optics is insufficient to find the necessary proportions, so we approach this problem using the full classical theory. We only need the (reasonable) assumption that the electric and magnetic field components lying on the boundary must be the same in both mediums.

Consider a wave traveling downward at an angle with respect to the y-axis. Let it be S polarized, i.e. the electric field is orthogonal to the x-y plane and the magnetic points downward at an angle with respect to both the x and y axes. Working only with the field magnitudes, the assumed continuity imposes the conditions

$$\begin{aligned} E_t &= E_i + E_r \\ -B_t \cos\theta_t &= -B_i \cos\theta_i + B_r \cos\theta_r \end{aligned}$$

In reality we should be matching Hs instead of Bs, but we make the approximation that the materials are non-magnetic. In the above, θ_i represents the angle the incident ray makes with the normal, θ_r is the same for the reflected ray, and θ_t marks the angle the transmitted ray makes with the normal. Obviously, $\theta_i = \theta_r$.

Since we're only dealing with amplitudes, we can make the substitutions $B = E/v = nE/c$ into the second equation and eliminate E_t to get

$$n_2(E_i + E_r)\cos\theta_t = n_1(E_i - E_r)\cos\theta_i$$

Collecting E_is and E_rs gives

$$E_i(n_1 \cos\theta_i - n_2 \cos\theta_t) = E_r(n_2 \cos\theta_t + n_1 \cos\theta_i)$$

yielding a reflection coefficient of

$$R_S = \frac{E_r^2}{E_i^2} = \Big(\frac{n_1 \cos\theta_i - n_2 \cos\theta_t}{n_1 \cos\theta_i + n_2 \cos\theta_t}\Big)^2$$

where we have squared the electric fields because we are concerned with the transmission of power, which is proportional to E^2. This gives a transmission coefficient of $T_S = 1 - R_S$.

Going through the same process with P polarized light (where the magnetic field is orthogonal to the plane of reflection) yields

$$R_P = \Big(\frac{n_1 \cos\theta_t - n_2 \cos\theta_i}{n_1 \cos\theta_t + n_2 \cos\theta_i}\Big)^2$$

where, again, $T_P = 1 - R_P$. Since we deal with unpolarized light in almost all instances within computer graphics, we can take

$$R_{\text{eff}} = \frac{R_S + R_P}{2}, \qquad T_{\text{eff}} = 1 - R_{\text{eff}}$$

As a final form for the effective reflection coefficient, we can express θ_t in terms of θ_i by using Snell's law. This results in an ugly—and computationally inefficient—equation.

Christophe Schlick suggested the approximation

$$R_{\text{eff}}(\theta) = R_0 + (1 - R_0)(1 - \cos\theta_i)^5$$

where

$$R_0 = \Big(\frac{n_1 - n_2}{n_1 + n_2}\Big)^2$$

and this is the approximaton we use in the accompanying ray tracer.

Gamma Correction and Desaturation

In the conversion from radiance to RGB, gamma correction must be applied. In short, modern displays use phosphor, which acts a nonlinear light source. As a standard, manufacturers make sure that the phosphor is implemented in such a way that the output-input relationship is simply

$$L_o = L_i^{2.2}$$

From this, we see that we should scale our radiance values by the 2.2 root before saving. This is often approximated by $\sqrt{L}$, but we use the exact value in the accompanying ray tracer.

Desaturation fixes a byproduct of digital image formats: since pixels can only take on a finite set of values, an extremely bright light source will change hues in the image (once the largest value hits 255, the increase among all three values is no longer proportional). We can workaround this by interpolating between the computed radiance and pure white using the peak of the computed spectra as the interpolating variable.