Prisms and Rainbows: a Dispersion Model for Computer Graphics

F. Kenton Musgrave

Department of Mathematics
Yale University
Box 2155 Yale Station
New Haven, Connecticut 06520

Abstract

Dispersion is the spreading of refracted light into its component colors or spectrum. A model of refraction including dispersion is developed using the techniques of distributed ray tracing. Two models of the rainbow, one empirical or impressionistic, the other purely physical, are developed using the results of the dispersion model. The problem of representing the spectrum of monochromatic colors using the rgb primaries of the graphics monitor is addressed.

KEYWORDS: Dispersion, refraction, rainbow, stochastic sampling, distributed ray tracing, spectrum, color gamut.

1. INTRODUCTION

Treatment of refraction in computer graphics has generally lacked *dispersion*, or the spreading of refracted light into its component colors or *spectrum*. While convincing simulations of transparent objects can be had without taking dispersion into account, the inclusion of dispersion makes available additional realism and beauty. We will present a dispersion model, within the *ray tracing* paradigm, and develop a physical model of the rainbow based on that dispersion model.

Modelling of dispersion entails the solution of at least two distinct problems: the integration and reconstruction of the *power spectrum* of light by frequency, and the display of the spectrum of *monochromatic* colors on a standard graphics display device. The first problem may be treated as another aspect of the *distributed ray tracing* model of Cook[4] et. al. or as an enhancement to the *rendering equation* of Kajiya.[11] The problem of reproducing monochromatic colors is in the realm of color science[27] and an approximate solution can be had through the use of *metamers*, though this problem remains an open area of research.

Perhaps the most spectacular example of dispersion at work in nature is the rainbow. The arc of the rainbow is a result of the geometry of the reflection and refraction of light in raindrops; the wonderful colors of the rainbow are the result of dispersion of sunlight in refraction through water. With a working dispersion model and some geometric optics, one can produce an efficient rainbow model for use in ray-raced and Z-buffered rendering schemes. We will present two rainbow models, one impressionistic or empirical[8] and another purely physical and therefore, quite true to nature.

2. PROBLEM STATEMENT

The Cook-Torrance[3] shading model takes into account the frequency of light waves in reflection from surfaces as a function of the index of refraction. What has been missing from the generally available literature is a model of refraction which takes into account the frequency of light. Such a dispersion model has been called for in previous research.[11,13] Some dispersion models have apparently been developed, but not published.[9,25] Thomas[23] published a brief description of a dispersion model, but did not develop atmospheric rainbows; unfortunately, Thomas' article remains obscure. The work presented here was developed independently of Thomas, and differs in most important respects.

The model of dispersion developed here is an extension of distributed ray tracing[24] and thereby uses the Monte Carlo integration techniques of Cook.[6] Integration of a continuous function by a finite number of *point samples* can lead to two types of *aliasing*, that of the frequency content of the signal being sampled and that introduced in the reconstruction of the signal from the samples. It is important to note that we are not concerned with the former type of aliasing, which is the result of sampling the signal at a rate below the Nyquist limit. Color *metamerism* generally obviates the need for very accurate reproduction of the exact curve of the power spectrum; nuances of the power distribution are important only in the interaction of light with attenuating media and reflecting surfaces and can safely be ignored in our model. What **is** important is our reconstruction of the spectrum from the point samples taken. As our approximation of the integral of the power spectrum will be a set of discrete samples, our reconstruction will be prone to appearing as a set of discrete, overlapping colors. This situation is analogous to that of *temporal aliasing*, where a moving ball may be sampled (imaged) at several points in time in an attempt to get motion blur and, upon reconstruction, appear as several overlapping, translucent circles.

In the case of dispersion, if we were to view a white disk on a black background through a prism, we might see several overlapping disks of different colors. We call this effect *spectral aliasing*, and use the *jittering* technique of stochastic sampling to defeat it. Jittering is random placement of the actual sample points within fixed sample intervals, which intervals may themselves be regularly spaced. Jittering adds noise to the image and turns the distinct overlapping images into a speckled blur, which looks a bit like spray paint.

The advantage of this noisy reconstruction of the image is that the human visual system tends to blur the noise together into a smooth continuum, whereas it actually enhances the sharp edges in the non-noisy images for a most displeasing effect. Such sharp discontinuities in intensity or color, or the rate of change thereof, manifest the phenomenon known as *Mach banding*. Mach bands are an artifact of the edge-enhancement caused by *lateral inhibition* in the retina.[7] When constructing and sampling our representation of the spectrum we must be aware of the potential for trouble with color Mach banding. The practical significance of this problem will be addressed in section 4.1.

Whatever colors we choose for representation, we will fail to accurately reproduce the spectrum. The graphics monitor has three *primary* colors with which to work, none of which is fully *saturated*. Even if we have three fully saturated or monochromatic primaries (as are available with laser raster projection systems), all other monochromatic colors can only be approximated, with varying degrees of *desaturation*. Our task, then, is to represent the entire visible spectrum of monochromatic colors as best we can, using three desaturated primaries and avoiding Mach bands. Furthermore, the sum of the samples chosen to represent the spectrum must, at full intensity, be the value of full-intensity white. If not, image samples involving dispersion will be tinted and/or shifted in intensity.

Given a working model of dispersion and an acceptable representation of the spectrum, one looks for applications. One striking application is a physical model of the rainbow. Rainbows are the result of the interaction of sunlight with very large numbers of raindrops in the atmosphere. The sheer number of particles (raindrops) involved, multiplied by the number of samples required to integrate the spectrum, makes a direct simulation of nature quite impractical. By modelling of the interaction of light with a single ideal raindrop, we can acquire a table of data which represents the situation in nature. This table may be used subsequently in the rendering process to replicate the effects of a rainbow in nature, with very good computational efficiency. We will describe such an approach in Section 4.3.

3. PREVIOUS WORK

3.1. Physics of Refraction

Refraction is an effect of the differing speed of light in dissimilar materials. The speed of light in a material determines its *optical density* which, surprisingly, is not exactly proportional to its mass density. As light slows down upon entering a medium of greater optical density, the wave trains are compressed. Thus, while frequency is preserved, wavelength is not. (It thereby behooves one to be careful **not** to use "frequency" and "wavelength" interchangeably when discussing refraction and dispersion.)

The *angle of refraction*, or the angle of the change in path for light, was related mathematically to the net change in *index of refraction* by Willebrord Snell in 1621 by Snell's Law:[1, 19]

$$\eta_1 \sin\theta_i = \eta_2 \sin\theta_t \tag{1}$$

where η_1 and η_2 are the indices of refraction of the two transmissive media, θ_i is the angle of incidence and θ_t is the angle of transmission. As the refractive index η is a function of the frequency of the light ray, the angle of refraction is also a function of frequency. Thus arises dispersion.

3.2. Physics of Dispersion

The proportion of change of index of refraction with frequency in a material is termed *dispersive power.* The dispersive power w of a material is defined as the ratio of the dispersion between the F and C Fraunhofer lines* to the mean deviation, i.e., the deviation for the D Fraunhofer line.[20, 21, 26] Thus

$$w = \frac{(\eta_F - \eta_C)}{(\eta_D - 1)} \qquad (2)$$

where η_F, η_C, and η_D are the refractive indices of the material at the frequencies of the F, C, and D Fraunhofer lines, respectively.

Just as optical density is independent of mass density, dispersive power is independent of optical density. The reason is that dispersion is modulated by *absorption bands* in materials, not by optical density.

Figure 3.1 The dispersion curve at an absorption band.

Note also that the plot of refractive index vs. frequency is not perfectly straight, but curved. This is an important factor in the development of a model of dispersion.

There have been many attempts to formulate a quantitative relation of refractive index η to frequency or wavelength λ, none entirely successful. The best known and most general is that of Sellmeier:[1]

$$\eta^2 = 1 + \sum \frac{b\lambda^2}{c^2 - \lambda^2} \qquad (3)$$

where b is a constant characteristic of the material, c is

* The Fraunhofer lines are *emission lines* of hydrogen. They represent monochromatic light at various visible wavelengths: the C line is at 656.3 nm (red), D is at 589.3 nm (yellow), and F is at 486.1 nm (violet).

an idealized absorption wavelength of the material (corresponding to a spectral *absorption band*) where the index of refraction is infinite, and the summation is over all absorption bands in the material. Simpler equations which are suitable for limited extents within the spectrum are:[1]

$$\eta = \frac{a}{\lambda^0} + \frac{b}{\lambda^2} + \frac{c}{\lambda^4} + \cdots \qquad \text{(Cauchy)}$$

$$\eta = 1 + \frac{b}{(c - \lambda)} \qquad \text{(Hartmann)}$$

$$\eta = a + \frac{b}{\lambda} + \frac{c}{\lambda^{\frac{7}{2}}} \qquad \text{(Conrady)}$$

$$\eta = a + b\lambda^2 + cL + dL^2 \qquad \text{(Hertzberger)}$$

where $L = (\lambda^2 - 0.028)^{-1}$, and a, b, c, and d are constants. These equations are all nonlinear, and values of the constants for various materials are not easily found in the literature. This will be a consideration in our development of a dispersion model.

3.3. Rainbows

Rene Descartes worked out the first scientifically accurate model of the rainbow in 1637.[10, 14] To do this, he assumed the raindrops to be spherical and traced rays through a circular, two dimensional representation - proof that ray tracing is hardly a new technique! Descartes' simulation is illustrated in Figure 3.2.

With his simulation, Descartes was able to accurately explain the angular size and position of the primary rainbow arc and some of the supernumerary arcs. (The supernumerary arcs which sometimes appear immediately inside of the primary rainbow arc are due to *diffraction* effects arising from the wave nature of light, and thus cannot be modelled using the geometric optics of a particle transport ray tracing paradigm. For more on this topic, see Nussenzvieg.[17]) Interestingly, an explanation for the color in the rainbow had to await Newton's discovery of dispersion some decades later. Aside from the supernumerary arcs inside the primary rainbow arc, Descartes' raindrop remains an accurate and sufficient model of the rainbow.

To recreate Descartes' simulation, we trace rays into the raindrop from the optical axis (ray 1 in Figure 3.2) to

Figure 3.2 Descartes' raindrop.

the edge of the circle. This corresponds to a range of zero to one for the *impact parameter*; the value of this impact parameter uniquely determines the path of the ray through the raindrop. Upon impinging the the raindrop, the ray is refracted, reflected once for the primary arc or twice for the secondary arc, and refracted again upon exiting the drop. Arcs formed by higher-order internal reflections are deemed unimportant as they are too dim and/or appear close to the sun in the sky, and are therefore not visible.

Note that all rays with an impact parameter greater than or less than that of ray 7 in Figure 3.2, the *Descartes ray*, emerge at an angle closer to the optical axis than that ray. Thus the Descartes ray marks a point of inflection in the change of emergence angle with impact parameter, and there is a concentration of light energy being returned at this angle, which is approximately 42 degrees. This gives us a bright feature 42 degrees from the optical axis; it is dispersion which spreads the bright feature into the spectrum of colors. Note also that the fact that all rays which are reflected exactly once inside the raindrop emerge at 42 degrees or less, makes the sky appear lighter inside of the primary arc of the rainbow. Rays reflected exactly twice inside the raindrop emerge with a peak power at approximately 52 degrees, with the excess light emerging at greater angles. Thus the secondary arc appears at about 52 degrees; between the two arcs is a zone of darkness known as Alexander's band.

* The extinction coefficient is a physical quantity specific to each material[22] which varies with frequency. The specific values of this coefficient are often unknown for a given material, and it is generally set to 0, for the purposes of computer graphics lighting models.

To perform an accurate simulation of energy transfer in Descartes' raindrop, the Fresnel equation should be used to modulate the quantities of reflected and refracted energy. With an *extinction coefficient** of 0, the Fresnel equation for reflection can be written:[2]

$$r_{pa} = \frac{\eta_2 \cos\theta_i - \eta_1 \cos\theta_t}{\eta_2 \cos\theta_i + \eta_1 \cos\theta_t} \tag{4}$$

$$r_{pe} = \frac{\eta_1 \cos\theta_i - \eta_2 \cos\theta_t}{\eta_1 \cos\theta_i + \eta_2 \cos\theta_t} \tag{5}$$

$$R = \frac{r_{pa}^2 + r_{pe}^2}{2} \tag{6}$$

where r_{pa} is the reflection coefficient for the component of light which is polarized parallel to the surface, r_{pe} is the reflection coefficient for the component polarized perpendicular to the surface, η_1 and η_2 are the refractive indices of the two materials, θ_i is the angle of incidence, θ_t is the angle of refraction, and R is the total reflectivity. Light not reflected is refracted in quantity $1-R$.

The rainbow phenomenon exists as a cone in space which is unique for each point of view (and indeed for each eye of the individual observer); Figure 3.3 is intended to illuminate this. Inspect it carefully for the following argument. Since the geometry of reflection and refraction as discussed above gives us a spectrum appearing at an angle the same as that of the Descartes ray from the straight back direction to the light source, we would expect to see that spectrum in all (sunlit) raindrops viewed from that angle. The sun's rays can be assumed

Figure 3.3 The cone of a rainbow.

to be parallel, thus this effect appears to the observer as a circle of angular radius 42 degrees, since the observer is, by definition, at the apex of the cone. Naturally occurring rainbows actually constitute a cone of half-angle 42 degrees around the *antisolar point* and have an angular width of approximately 2 degrees. The secondary arc appears at a half-angle of 52 degrees.

3.4. Computer Graphics

As mentioned above, the Cook-Torrance shading model relates reflection to index of refraction and frequency through the Fresnel equation.[22] A model of refraction relating index of refraction to frequency has been developed by Thomas[23] and more recently by the author;[15,16] that work is extended here to include a physical model of the rainbow.

The problem of integration and reconstruction using point samples has been addressed by Cook[6] in his discussion of the distributed ray tracing model.[4] The dispersion model developed by the author is a straightforward application of Cook's techniques, as an extension to the repertoire of effects available through distributed ray tracing.

A model of atmospheric rainbows has been alluded to in the literature,[5] but not presented in detail. A physical model of the rainbow requires a fair amount of development work. Fortunately, the development work being done, the results are easy to include as an added feature in a rendering program.

4. SOLUTION

4.1. Sampling in the Frequency Domain of Light

To model dispersion, we must integrate the power spectrum of light at each sample point in the image where there occurs dispersive refraction, such as on the surface of a glass prism. The integral of the power spectrum can be expressed

$$I_T = \int_{380}^{800} I(\lambda)\ d\lambda \tag{7}$$

where I_T is the total illuminance at the given point in space and $I(\lambda)$ is the illuminance at wavelength λ at that point. As we need only integrate the power spectrum of *transmitted* light at dispersive surfaces, since only transmitted or refracted light is dispersed, the integral we are interested in can be stated

$$I_t = \int_{380}^{800} T(\lambda)\ d\lambda \tag{8}$$

where I_t is now the illuminance by transmitted light at a point in space on the boundary of a change in refractive index, and $T(\lambda)$ is the illuminance by the transmitted light at wavelength λ.

As previously stated, we will approximate this integral using a set of point samples. We perform *stochastic antialiasing* of our integral by *jittering*[6] the samples. If a sample f at frequency λ represents the power in the spectrum over an interval of width Δf, the jittering consists of adding a random offset $\Delta f\ (X - 1/2)$ where X is a random variable of uniform distribution in the range [0..1]. The net effect is to randomly place the sample f somewhere within the interval $\lambda - \Delta f/2$ to $\lambda + \Delta f/2$.

The fact that we take point samples in the frequency continuum of light implies that we are also taking point samples of the continuum of the dispersion curve, as index of refraction is a function of frequency. Thus we face the choice of whether to jitter the frequency (and therefore the color) of the rays or the refractive index of the material, or both. Given that the the jittered sample at frequency f needs to be translated into $R(f)$, the value of the refractive index function R at f, we will prefer to jitter a linear function R over a nonlinear function for reasons of computational efficiency, as linear interpolation is in general quicker to evaluate than nonlinear interpolation.

This may motivate us to contrive piecewise linear approximations to the spectrum and the dispersion curve. It is unlikely that the viewer of the final image will be able to discriminate between a physically accurate nonlinear model and a computationally efficient linear approximation; furthermore, since the dispersion curve is specific to a given material, to be true to nature one would need to tabulate data for every distinct material to be rendered. We therefore employ a (one-piece) linear approximation to the dispersion curve for our rendering dispersion model.

The refractive index and dispersive power for surfaces can be input parameters. Thus one can specify a polygon with an associated refractive index of, for example, 4.2 and a dispersive power of perhaps 0.5, both of which are outlandish in terms of the "real" world, but viable within our model. It is interesting to create situations and materials which cannot exist in our everyday experience; this is part of the power of computer graphics.

The issue of which quantity to jitter, refractive index or color, or both, should be evaluated in the light of computational efficiency. The reason for jittering samples is to avoid spectral aliasing, however, it has been our experience that spectral aliasing is not a significant problem in any but deliberately pathological scenes. That is, the distinct overlapping images of different colors are simply not readily visible unless the dispersive power is unrealistically high. When jittering is deemed desirable, we jitter the frequency of the ray and derive, in a prerendering operation, a constant c_s for each refractive surface s in the scene:

$$c_s = \frac{w(\eta - 1) - \eta}{0.76} \tag{9}$$

where w is the dispersive power, η is the refractive index at the far red end of the spectrum, 0.76 is the proportion of the spectrum that lies between the C and F Fraunhofer lines. This constant c_s when multiplied by the the frequency of a sample gives the refractive index at that frequency, for use in calculations of propagation of refracted light. (Note that this assumes that frequency is specified in the range [0..1].) Thus the cost of jittering is reduced to one floating point multiplication per surface encountered, plus the negligible preprocessing cost of evaluating c_s for each relevant object in the scene and the cost of interpolating the color of the final sample.

4.2. Representing the Spectrum

To reproduce the spectrum, we must simulate the entire gamut of monochromatic colors using only the three desaturated primaries of the graphics monitor. Furthermore, the integral of each of the red, green, and blue curves of our simulated spectrum must be unity, or the reconstruction of an image from our samples will be tinted, darkened, or overdriven. We refer to this as the *summing to white* criterion.

As we work within the *rgb* color space, we should restate equation (9) in terms of the rgb vectors:

$$I_{t_R} = \int_{380}^{800} R(\lambda)\, T(\lambda)\; d\lambda \tag{10}$$

$$I_{t_G} = \int_{380}^{800} G(\lambda)\, T(\lambda)\; d\lambda \tag{11}$$

$$I_{t_B} = \int_{380}^{800} B(\lambda)\, T(\lambda)\; d\lambda \tag{12}$$

where $R(\lambda)$, $G(\lambda)$, and $B(\lambda)$ are the values of the R, G, and B *tristimulus* functions for the metameric color used to represent the color of monochromatic light of wavelength λ. When sampling at a particular frequency then, we are actually taking three (red, green, and blue) samples of $T(\lambda)$. The distribution of the samples should be tailored to the shape of the tristimulus curves used in the representation of the spectrum, with care taken to assure that

$$\sum_{i=1}^{n} R(\lambda_i) = \sum_{i=1}^{n} G(\lambda_i) = \sum_{i=1}^{n} B(\lambda_i) \tag{13}$$

where λ_i is the wavelength of the i^{th} sample, and $R(\lambda_i)$, $G(\lambda_i)$, and $B(\lambda_i)$ are the red, green, and blue values, respectively, of sample λ_i. This equality is necessary in order to have the samples (at their maximum intensity values) sum to white in the rgb color space of the graphics monitor.

4.2.1. Linear Spectrum Model

A simple representation of the spectrum, given these constraints, is shown in Figure 4.1.

Figure 4.1. The rgb curves of the linear spectral representation.

This model has the advantage of being piecewise linear, for fast interpolation of color, and it provides a reasonably good perceptual representation of the spectrum. It has the disadvantage of using a significant portion of the power available to the red primary, in the approximation of violet with magenta. Violet is of higher frequency than is available with an rgb monitor and therefore cannot be directly reproduced; magenta is a visually acceptable substitute. A problem with the magenta representation of violet is that edges which are blurred by dispersion such that they should appear with the color sequence yellow-orange-red-black, actually appear greenish-yellow-red-black. This is because in a white-to-black transition of this sort, the first color to be subtracted out from the sum is violet. When violet is represented as a sum of equal quantities of red and blue, the subtraction of violet leaves a surplus of green. This is a subtle effect, and escapes the notice of most viewers.

Another potential drawback of this representation of the spectrum is the pronounced discontinuities in the first derivative of the rgb curves. While this has the potential for causing color mach banding, such an effect has only been observed in deliberately pathological scenes. Yet another problem found is that the red band in the spectrum appears too narrow, again because some of the red energy is used to display violet. The final problem is that the rolloff of red and violet to black is too steep and short; the entire curve bears no resemblance to the response curve of the human visual system.

Despite the above drawbacks, we have found this to be a viable representation of the visible spectrum.

We sample the representation of the spectrum at 13 intervals centered on the vertical lines in Figure 4.1. This provides a good basis for reconstruction of the spectrum and preserves the summing to white property. However, when jittering we encounter the problem that the samples may longer sum to white. The noise added by uncorrelated jittering of the 13 samples will generally skew the sum; in practice this appears as a faint colorful noise, faint enough to not be objectionable or even usually noticeable. (This problem could be defeated by correlating the jittering of the 13 samples, but this is computationally expensive.) Furthermore, about half the time the sum of jittered samples of a full intensity white point will exceed unity. If the sum is not clamped to unity at the high end, overflow will occur and the color of the summed samples is likely to wrap around to black. This problem is defeated by clamping the sum, at minimal computational cost.

4.2.2. Empirical Spectrum Model

A more rigorous approach to the construction of the representation of the spectrum is currently under development. This approach involves taking the *xyz* coordinates of the monochromatic spectral colors and performing the appropriate linear transformation into rgb values. Construction of the transformation matrix requires information about the chromaticity coordinates of the specific monitor on which the spectrum is to be displayed.[19] We use as input the xyz coordinates of monochromatic colors weighted by the spectral radiant power distribution of the CIE standard illuminant B, which is designed to emulate direct sunlight (the light source for rainbows). The following graphs are piecewise linear between samples taken at 10nm (nanometer) intervals from 380 to 770 nm.[27]

Figure 4.2. The rgb curves of the empirical spectral representation.

Figure 4.3. Summed rgb values, with and without negative values.

As our rgb primaries are not fully saturated, we expect that at all points in the spectrum at least one of the rgb values will be negative. This is indeed what we see in the curves of Figure 4.2. The sum of these curves, with negative values included and with negative values clamped to zero, are shown in Figure 4.3. A more accurate approximation to the spectrum, without negative values, could be attained be limiting the xyz input values to the color gamut of the monitor.

Note that the curves in Figure 4.3 have a local minimum in the cyan area of the spectrum. These curves do not give an acceptable representation of the spectrum on a monitor calibrated for perceptually linear contrast response; the cyan and yellow colors appear far too dark. When adjusted with a *gamma correction* of 2.5 to 3.0, however, the zero-clamped curve gives a good representation of the spectrum. Note also that the area under the zero clamped curves should be normalized to meet the summing to white requirement.

4.3. Rainbow Models

4.3.1. Impressionistic Rainbow Model

We have developed two models of the rainbow, one very simple and impressionistic or empirical, the other comparatively complex and purely physical. The former model entails using the 13 colors of our samples of the linear spectrum model to create 13 different colors of fog which compose a rainbow. The fog function is simply an asymptotic replacement of some percentage r of the color value computed at the end of the ray, with the color value of the fog, based on the distance that the ray has traveled:

$$r = e^{hd/t} \tag{14}$$

where h is a constant, d is the distance, and t is the *transmittance* constant; note that t has red, green, and blue components, usually equal. As that distance goes to infinity, the percentage of replacement goes to 100. The 13 colored fogs are invoked in concentric rings (cones, actually) around the *antisolar vector*, e.g., the vector from the light source to the eye point. This vector corresponds to the ray from the observer to the antisolar point in Figure 3.3. Each ring is a band of some angular width, at some angular offset from the antisolar vector. We construct the rainbow by taking the dot product of each ray traced, with the antisolar vector; this dot product gives us the cosine of the angle between the two vectors. This cosine is then used as an index into a table of the 13 colored fogs. The indexing function can be parameterized to vary the width and angular placement of the rainbow. The following C code segment implements this parameterized rainbow:

```
index = ( DOT(ray_direction, antisolar_ray)
          − rainbow_angle) * rainbow_width;
if (jitter_option)                                  if
        index += jitter(delta);
if ((index >= 0) && (index < FREQUENCIES))          if
        Fog = Rainbow[(int)index];
else Fog = NULL;
```

where "ray_direction" and "antisolar_ray" are vectors, "Fog" and "Rainbow[]" are pointers to structures for the "fog" type, and the other variables are floating point type. The constant "FREQUENCIES" is equal to 13; the function call "jitter(delta)" returns a random value of uniform distribution in the range [-delta/2..delta/2].

The jitter option turns a rainbow composed of concentric bands of color to a more attractive "fuzzy" rainbow. This jittered rainbow can look fairly realistic, particularly when supersampling is employed to soften the noise introduced by the jittering. Note that this scheme only jitters the index to the table of colored fogs, and not the color of the fog itself; an improvement would be to add such color jittering.

4.3.2. Physical Rainbow Model

The above approach is *ad hoc* and is not really based on a dispersion model, but it uses the spectral representation of our dispersion scheme. A more rigorous and complex approach, yielding a more realistic result, is to recreate Descartes' simulation using dispersion. We will have to integrate Descartes' raindrop over the visible frequencies of light; this entails ray tracing Descartes' raindrop at a variety of frequencies and summing the results. Clearly it is inefficient to ray trace Descartes' raindrop for every ray spawned in the process of rendering a scene; fortunately we can do much better than this. We need only perform the integration over frequency of Descartes' raindrop once, in a preprocessing step, to build a table of fogs similar to that used in the our simpler rainbow model. This table will need to have a relatively large number of entries, as a real rainbow is an illumination effect that covers most of the sky, though mostly to a very subtle degree. Thus we have entries for a large number of angular displacements, over a 180 degree range. (In practice, no fog might be required in the 10 degree interval of Alexander's band, as no light is returned there by refraction.)

The first step in implementation of the physical model is to generate an algorithm for ray tracing Descartes' raindrop. This means calculating the angle of emergence and energy attenuation factor for rays which are reflected once and twice inside the raindrop, as a function of the impact parameter. The angle of emergence of a given ray is determined by the geometric optics of reflection and refraction in a sphere, while the energy transfer is determined by the physics of reflection and refraction of light as it interacts with air/water boundaries.

The geometric optics of Descartes' raindrop are illustrated in Figure 4.4.

Figure 4.4 The geometric optics of Descartes' raindrop.

Note that we can take advantage of the equality of angles θ_1 and θ_2. Once this geometry is established, it is straightforward to program an algorithm to trace the required rays.

For the purposes of computer graphics, we are generally not concerned with the polarization of light, and generalizations of the Fresnel equation for non-polarized light are usually employed. For this simulation, however, we are more interested in physical veracity than computational efficiency, so we choose the formulation of the equation as it appears in equations 4, 5, and 6. Note that the orientation of polarization to the surface is preserved through reflections and refractions in a spherical raindrop.

Also in the interest of physical accuracy, we use a nonlinear approximation of the dispersion curve of water in our rainbow simulation. Using actual measurements of the refractive index of water at various frequencies[12] we derive constants a and b of Cauchy's equation for refractive index, getting $a = 1.3239$ and $b = 3116.3$. The first two elements of the Cauchy series

$$\eta = \frac{a}{\lambda^0} + \frac{b}{\lambda^2} = a + \frac{b}{\lambda^2} \tag{15}$$

give a good approximation to the dispersion curve of water with the derived values of a and b: over the wavelength range from 405 to 670 nanometers, the calculated values of η agree with measured values to within plus or minus 0.0001, or 0.8 of one percent. We use a refractive index of 1.0003 for air.

Our first implementation of the physical rainbow model uses samples taken at 13 fixed, evenly spaced frequencies or wavelengths. (We relax our rigor in the use of "frequency" and "wavelength" here, as the visible spectrum is usually specified by wavelengths of light in a vacuum.) We trace 50,000 rays per wavelength, over the range of impact parameters. For each wavelength sampled, the intensities of the emerging rays are summed by angle of emergence in a linear array of 1800 buckets. The intensities in each bucket are then multiplied by the rgb vector of the representative color for that wavelength and added to buckets of a similar array of rgb intensities by angle. After all wavelengths have been sampled, the results in the rgb array are normalized and inverted for use in the fog function. Unlike the *ad hoc* rainbow model, the fogs used are not themselves colored, but rather their transmittances, t in equation (14), are unequal in red, green, and blue. Thus the fogs have no intrinsic color, but red, green, and blue values at ray endpoints are replaced at independent rates per unit distance. This prevents unnecessary filtering by attenuation of colors behind the rainbow.

Our first approach evidences significant spectral aliasing. Spectral aliasing is accentuated in the rainbow model, as the bright feature at the Descartes ray is quite narrow and pronounced for a point light source, resulting in thin concentric rings of color in the rainbow. The rings are more widely spaced and therefore more evident in the violet end of the spectrum, as the dispersion curve is steeper at shorter wavelengths.

A second implementation employs spectral antialiasing. Again we sample at 13 distinct frequency intervals, but we jitter the samples within the intervals. This approach requires that we multiply the intensity of the ray by the interpolated rgb value for its specific frequency, and store that vector in the rgb array immediately, rather than using an intermediate storage array, as the colors of individual rays will vary. This has the effect of blurring and merging the rings produced by discrete sampling.

Again, the process described above yields the rainbow produced by a point light source, thus the rings of color produced by spectral aliasing are quite narrow and distinct. In nature rainbows are produced by the sun, which has an angular diameter of approximately one half of one degree. Convolution of the final rgb tables with a (one dimensional) kernel which represents the disk of the sun spreads each of the rings over one half a degree of angle. The kernel we use is five entries wide, corresponding to the fact that our fog samples are taken at $1/10^{th}$ degree intervals. Since the entire angular width of the rainbow is approximately two degrees, this blurs the rings together well enough to provide very good spectral antialiasing. If the area under the curve of the semicircular kernel is normalized, there will be no net change in the density of the fog tables after the convolution.

We employ another feature of our rainbow models. In nature, rainbows are rarely perfect arcs, in fact one most often sees only a portion of the full rainbow arc. Rainbows are modulated by two factors: shadows of the clouds from which the rain is falling, and the distribution of the falling rain itself. In an effort to make our rainbows look more natural, we modulate intensity of the rainbow with Perlin's[18] "Chaos()" texture. This is a *solid* or *procedural* texture which takes a vector as input and returns a stochastic scalar quantity with a $1/f^2$ power spectrum. The vector we pass to the texture is the ray direction; we use the scalar value returned to modulate the transmittance of the rainbow fogs. The frequency content of the Chaos() function can be parameterized for varying effects, and the texture can be scaled on a vertical or slanted axis to simulate sheets of falling rain.

5. CONCLUSION

A model of dispersive refraction within the distributed ray tracing paradigm has been implemented, with good subjective results. The problem of representing the spectrum of monochromatic colors within the rgb color space has been addressed, but not solved to final satisfaction; further work is called for here.

Physical and empirical/impressionistic models of the rainbow have been developed, using the above results. In contrast to the dispersion model, the rainbow models are relatively efficient to render, because of their table-lookup implementation. The rainbow models are suitable for Z-buffer rendering schemes, as well as ray tracing.

References

1. R. M. Besancon, *The Encyclopedia of Physics,* Van Nostrand Reinhold Company, New York, 1974.
2. Max Born and Emil Wolf, *Principles of Optics,* Pergamon Press, Oxford, 1980.
3. Robert L. Cook and Kenneth E. Torrance, ''A Reflectance Model for Computer Graphics,'' *Computer Graphics*, vol. 15, no. 3, pp. 307-316, August, 1981.
4. Robert L. Cook, Thomas Porter, and Loren Carpenter, ''Distributed Ray Tracing,'' *Computer Graphics*, vol. 18, no. 3, pp. 137-145, July, 1984.
5. Robert L. Cook, ''Shade Trees,'' *Computer Graphics*, vol. 18, no. 3, pp. 223-230, July, 1984.
6. Robert L. Cook, ''Stochastic Sampling in Computer Graphics,'' *ACM Trans. on Graphics*, vol. 5, no. 1, pp. 51-72, January, 1986.
7. T. N. Cornsweet, *Visual Perception,* Academic Press, New York, 1970.
8. Alain Fournier, ''Modelling Natural Phenomena,'' *SIGGRAPH Course Notes*, 1987.
9. Alain Fournier, *personal communications*, 1987.
10. R. Greenler, *Rainbows, Halos, and Glories,* Cambridge University Press, Cambridge, 1980.
11. James T. Kajiya, ''The Rendering Equation,'' *Computer Graphics*, vol. 20, no. 4, pp. 143-150, August, 1986.
12. G. W. C. Kaye and T. H. Laby, *Tables of Physical and Chemical Constants, 14th Edition,* p. 95, Longman Group Ltd., London, 1973.
13. Mark E. Lee, Richard A. Redner, and Samuel P. Uselton, ''Statistically Optimized Sampling for Distributed Ray Tracing,'' *Computer Graphics*, vol. 19, no. 3, pp. 61-67, July 1985.
14. M. Minnaert, *The Nature of Light and Colour in the Open Air,* Dover, New York, 1954.

15. F. Kenton Musgrave, ''A Realistic Model of Refraction for Computer Graphics,'' *Masters Thesis*, University of California at Santa Cruz, Santa Cruz, California, September, 1987.

16. F. Kenton Musgrave, ''A Realistic Model of Refraction for Computer Graphics,'' *Modelling and Simulation on Microcomputers 1988, conference proceedings*, pp. 37-43, Society for Computer Simulation, San Diego, Feb. 1988.

17. H. Moyses Nussenzveig, ''The Theory of the Rainbow,'' *Scientific American*, pp. 55-65, April, 1977.

18. Ken Perlin, ''An Image Synthesizer,'' *Computer Graphics*, vol. 19, no. 3, pp. 287-296, July, 1985.

19. D. F. Rogers, *Procedural Elements for Computer Graphics,* Mc Graw Hill, New York, 1985.

20. G. G. Slyusarev, *Aberration and Optical Design Theory,* Adam Hilger Ltd, Bristol, 1984.

21. James P. C. Southall, *Mirrors, Prisms, and Lenses,* MacMillan Company, New York, 1933.

22. Ephraim M. Sparrow and R. D. Cess, *Radiation Heat Transfer,* pp. 64-68, McGraw-Hill, New York, 1978.

23. Spencer W. Thomas, ''Dispersive Refraction in Ray Tracing,'' *Visual Computer*, vol. 2, no. 1, pp. 3-8, Springer International, January, 1986.

24. Turner Whitted, ''An Improved Illumination Model for Shaded Display,'' *CACM*, vol. 23, no. 6, pp. 343-349, June, 1980.

25. Turner Whitted, *personal communications*, 1987.

26. Robert W. Wood, *Physical Optics,* MacMillan Company, New York, 1911.

27. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Methods and Formulas,* Wiley-Interscience, New York, 1967.