This project came from MAA notes 29, Resources for Calculus Collection, Volume 3; Applications of Calculus, Editor Philip Straffin.

SOMEWHERE WITHIN THE RAINBOW

Author: Steven Janke, Colorado College, Colorado Springs, CO 80903

Calculus Needed: Derivatives of trigonometric functions, finding maxima and minima.

Area of Application: Optics, Meteorology.

The Problem: Explaining the Rainbow

My heart leaps up when I behold a rainbow in the sky...— Wordsworth

Whether it has been raining for just a few hours or for forty days and forty nights, if the sun appears and raindrops are still in the air, the world is treated to one of nature's most vivid spectacles, the rainbow. Imagine the mixture of fear and wonder that ancient people must have felt on seeing such a sight. These days the wonder is still there, but certainly the fear has lessened as we understand more of the physics involved in producing such a display of color in the sky. After setting aside the awe, our curiosity produces question after question. Why is the rainbow a circular arc? What determines how high it is in the sky? Why are there colors? Why is there a special order to the colors? Why is there occasionally a second rainbow above the first? Exactly where is the pot of gold?

Some Early History

The early explanations of the rainbow were understandably mythological in origin. The Greek goddess Iris was said to use the rainbow as a sign both of warning and of hope. The word "iridescent" probably comes from the connection to Iris. In African mythology, the rainbow was a large snake coming out to graze after the storm. Here again the event is both a sign of hope and one of fear, for the snake could gobble children that were too close to the ends of the bow. The ends do appear to touch the earth leading some to claim that great treasure was buried there. Yet in a less capitalistic vein, many American Indians saw the bow as a bridge anchored in this world and leading to the next.

In 578 B.C., Anaximenes, a Greek scholar, noted the relation between the rainbow and the sun. Rather than attributing the bow to celestial powers, he suggested that clouds bent the sun's light to produce the arc of colors. Aristotle used careful geometry, but faulty reflection laws, to establish the circular shape of the bow. Gradually, scholars began to see that both reflection and refraction of light had something to do with the rainbow phenomenon. In the fourteenth century, Theodoric of Frieburg and the Persian scholar Kamal al-Din al Farisi independently decided that drops of rain were the key. They looked closely at the way a globe of water affected light and were able to give correct qualitative explanations for the bow.

The rainbow has piqued the interest of many scholars in each of the last several centuries. The sixteenth century seems to have produced the most books on the subject, but few of them were of major importance. As you might expect, seventeenth century scholars like Kepler, Descartes, Fermat, and Newton all made significant contributions to the study of the rainbow. Even today, physicists continue to tidy up the theory. Understanding the rainbow is so tied with understanding the nature of light that until theories of light are complete, there will be open questions about the rainbow.

Reflection

Light from the sun, refracted and reflected by water droplets in the atmosphere, forms the rainbow, so the first step in explaining the phenomenon is to understand how light is bent by various substances. In 1657, the extraordinary mathematician Pierre de Fermat turned his attention to the bending of light and proved the main results by postulating a simple principle. Fermat suggested that in traveling from point P to Q, light follows a path which minimizes the total travel time.

Fermat's Principle. *Light follows a path which minimizes the total travel time.*

Consider first the reflection of light. It helps when discussing geometric problems with light to imagine that light travels along rays. So suppose we have a source of light rays at point P in Figure 1. Imagine that we detect one of the rays passing through point Q after reflecting off a surface. At what point R does the ray reflect off the surface?

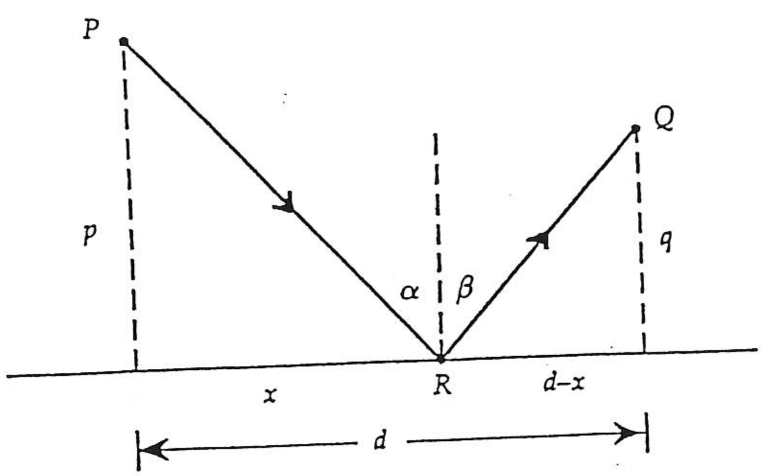

Figure 1.

Fermat's principle claims the ray follows a path that minimizes the time necessary to travel from P to Q while reflecting off the surface. Assuming the speed of light in our example is constant, the point R should be positioned so the path PRQ has minimum length. Considering the triangles in the figure, we get the following expression for the path length as a function of x:

$$L(x) = \sqrt{p^2 + x^2} + \sqrt{q^2 + (d - x)^2}.$$

To find the minimum path length, we find the derivative $L'(x)$ and set it equal to zero.

$$\begin{aligned} L'(x) &= \frac{1}{2}(p^2 + x^2)^{-\frac{1}{2}}(2x) + \frac{1}{2}(q^2 + (d - x)^2)^{-\frac{1}{2}} \cdot 2(d - x) \cdot (-1) \\ &= \frac{x}{\sqrt{p^2 + x^2}} - \frac{(d - x)}{\sqrt{q^2 + (d - x)^2}} = 0, \\ \text{so that } \quad \frac{x}{\sqrt{p^2 + x^2}} &= \frac{(d - x)}{\sqrt{q^2 + (d - x)^2}}. \end{aligned}$$

Now referring again to the figure, $\sin\alpha = \dfrac{x}{\sqrt{p^2 + x^2}}$. Similarly, $\sin\beta = \dfrac{d - x}{\sqrt{q^2 + (d - x)^2}}$. So $L'(x) = 0$ when $\sin\alpha = \sin\beta$. We should verify that this is a minimum by taking the second derivative (see the exercises). Rather than actually solve for the distance x, it is more useful to note that the minimum occurs when the sines of the two angles are the same. Since the angles are both between 0 and $\pi/2$, we conclude that the two angles are equal. For convenience call α the angle of incidence and β the angle of reflection.

Law of Reflection. *For reflection, the angle of incidence is equal to the angle of reflection.*

Note that we have deduced the Law of Reflection from Fermat's principle of least time. We have not proved Fermat's principle, but it does make sense in light of other results in physics. And in fact, careful experiments have concluded that the Law of Reflection does hold.

Exercises

1. Determine the value of x that minimizes $L(x)$ in the derivation of the Law of Reflection.
2. Compute the second derivative and use it to show that we indeed found a minimum for $L(x)$.

Refraction

When dealing with reflection, we assumed that the light rays were traveling only in air and therefore maintained a constant speed. However, to attack the rainbow questions, we need to also understand what happens when light travels through water. It turns out that the speed of light in water is less than the speed in air. Our derivation of the reflection law would be identical for a mirror and light source submerged in water since the speed of light would again be constant, but what happens if part of the light's path is in water and part is in air?

Figure 2 shows a new setup where point P is again a source of light rays in air. Now, however, point Q is in water. We are interested in the path of a light ray that leaves P and passes through Q. It crosses the air/water interface at the point R. The angle the path PR makes with the line perpendicular to the water's surface is called the angle of incidence and is represented by α. The corresponding angle between the path RQ and the perpendicular is called the angle of refraction and is represented by β. Fermat's principle claims that the point R is positioned so as to make the total time of travel a minimum. Since the speed changes when the light crosses into water, we need to consider both speeds in our analysis.

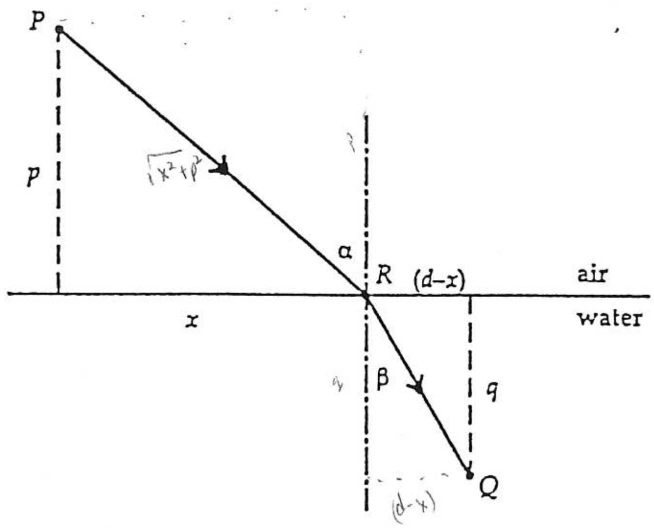

Figure 2.

Let c_a be the speed of light in air and let c_w be the speed of light in water. Remembering that time is distance divided by speed, we calculate that the light ray spends $\frac{\sqrt{p^2+x^2}}{c_a}$ units of time traveling from P to R and $\frac{\sqrt{q^2+(d-x)^2}}{c_w}$ units of time traveling from R to Q. Thus the total time is

$$T(x) = \frac{\sqrt{p^2+x^2}}{c_a} + \frac{\sqrt{q^2+(d-x)^2}}{c_w}.$$

Again to find the minimum, we take the derivative of $T(x)$:

$$\begin{aligned} T'(x) &= \frac{1}{c_a} \cdot \frac{x}{\sqrt{p^2+x^2}} - \frac{1}{c_w} \cdot \frac{d-x}{\sqrt{q^2+(d-x)^2}} \\ &= \frac{\sin\alpha}{c_a} - \frac{\sin\beta}{c_w}. \end{aligned}$$

Setting $T'(x) = 0$ gives

$$\frac{\sin\alpha}{c_a} = \frac{\sin\beta}{c_w}, \quad \text{or} \quad \frac{\sin\alpha}{\sin\beta} = \frac{c_a}{c_w}.$$

In other words the ratio of the sines is a constant. Note again that in order to verify that we have found the minimum we should take the second derivative.

This constant c_a/c_w is the ratio of the speed of light in air to the speed of light in water. In order to calculate it, tables have been compiled that give the ratio of the speed of light in a vacuum to the speed of light in various media. For example, the ratio of the speed in a vacuum to the speed in water is about 1.33 and is called the *index of refraction* for water. The index of refraction for air is very close to 1 so the ratio c_a/c_w is very close to 1.33.

Law of Refraction. *The ratio of the sine of the angle of incidence to the sine of the angle of refraction is a constant.*

In our derivation, there was no dependence on direction, so our result would be the same if we assumed that the source of light was at Q instead of at P. With this observation we notice that if light travels from one medium to one of higher refractive index, the light ray bends toward the perpendicular to the surface between the media. (This perpendicular is often called the normal.) When light travels from one medium to one of lower refractive index, the ray is bent away from the normal.

Fermat supplied the principle from which we mathematically deduce the Law of Refraction, but it was earlier, in 1621, that a Dutch scientist Willebrord Snell experimentally discovered the result. Today the Law of Refraction is often called Snell's law.

Exercise

3. Verify that we found a minimum for $T(x)$ in the derivation of the Law of Refraction.

The Rainbow Angle

Rainbows form when raindrops both reflect and refract light from the sun. When light traveling through the air strikes a drop, some of the light is reflected and some is refracted as it enters the drop. Part of the light inside the drop is reflected when it strikes the other side of the drop and part is refracted as it again passes into the air. In general, when light travels from one medium to another, part of the light is reflected at the interface and part continues into the second medium where it is refracted. To understand how the rainbow forms, we need to keep track of the reflections and refractions caused by a drop of rain.

The shape of a raindrop depends on several factors, but for a good approximation, it is fairly safe to assume that it is spherical. Look then at Figure 3. Here we see the cross-section of a drop as a light ray enters it at point A. Some of the light ray will be reflected, but the figure shows the part that enters the drop. The Law of Refraction says that this ray will be bent toward the normal since the refractive index of water is larger than that of air. From geometry, we know that the tangent to the circle at point A is perpendicular to the radius of the circle through A. Hence, the radius through A is the normal at A. In the figure, α is the angle of incidence and β is the angle of refraction.

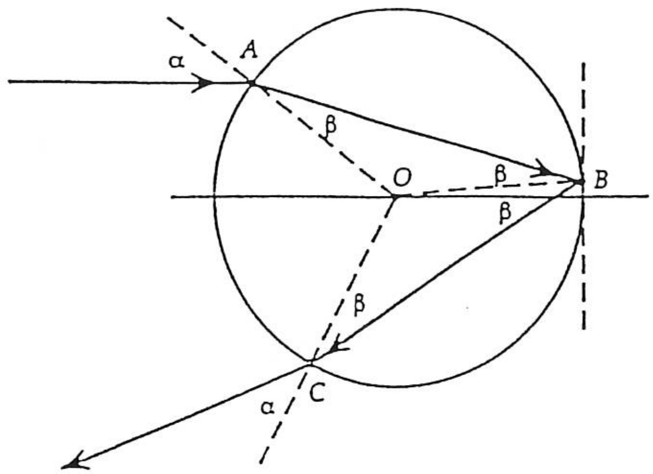

Figure 3.

The ray continues through the drop and strikes the other side at point B. Here again, part of the ray is reflected and part continues into the air where it is refracted. In the figure, we follow the reflected part. At B, the ray is reflected so that the angle of incidence equals the angle of reflection. Here the angle of incidence is the angle between the ray and the tangent at B. Notice that this implies that angle ABO equals angle OBC. When the ray hits the drop's surface at C, part is reflected, but let's follow the part that enters the air and is refracted. Since the ray moves into a medium of lower refractive index, it is bent away from the normal.

Figure 3 traces parts of one particular ray that strikes the drop. At each interface, another part is either refracted or reflected, and consequently there are many paths a ray could take in interacting with the drop. In fact, you can imagine a ray that enters the drop and is repeatedly reflected around inside it. Since at each interface between air and water part of the ray is reflected and part refracted, when we choose to follow one part we are following a ray that has less intensity than the original ray. Each time an interface is hit, the light intensity decreases. We are therefore interested in rays that strike the interface only a few times, for this will be the brightest light.

Again looking at the figure, a ray that strikes the drop at A and is simply reflected will be fairly bright, but as we will see, such a ray does not add to the essential features of the rainbow since it doesn't interact with the water. Similarly, a ray that hits at A and then travels through the drop to exit at B will also be fairly bright, but we would have to be on the righthand side of the drop to see this light. Rainbows are formed when the sun is behind us and light from it is reflected in various ways from the raindrops. So the ray drawn in the figure is the simplest ray involved in rainbow formation.

Notice that the point A could be anywhere on the left half of the circle. If it is on the upper half of the circle then the ray exits the drop in the lower half. We are interested in how much the ray is deflected once it exits the drop. For example, if the ray comes in the drop along the diameter of the circle, then the angle of incidence is zero and therefore the angle of refraction is zero. The ray will reflect off the back of the drop and exit the drop along the same diameter that it entered on. The total deflection would be 180 degrees in a clockwise direction. The ray drawn in the figure has been deflected by less than 180 degrees. As the point A moves on the circle, the deflection angle changes. So the angle of deflection is a function of the angle of incidence. If α is the angle of incidence, let $D(\alpha)$ represent the angle of deflection.

Because of the symmetry between the upper and lower halves of the the circle, we might as well focus only on those points A on the upper-left quarter of the circle. For these points, α varies from 0 to 90 degrees. To determine the total deflection, first consider how the ray is deflected at the point A. It is rotated clockwise by $\alpha - \beta$ degrees. At B, it is again rotated clockwise by $180 - 2\beta$ degrees. Finally at C the deflection is again $\alpha - \beta$ degrees. Hence

$$D(\alpha) = \alpha - \beta + 180 - 2\beta + \alpha - \beta = 180 + 2\alpha - 4\beta.$$

Notice that D is a function of both α and β. However, we know from the Law of Refraction that β can be expressed as a function of α. We will need to keep this in mind when we take the derivative.

Now $D(C) = 180$ and as α increases, $D(\alpha)$ at first decreases. But what is interesting is that D has a minimum. It only decreases so far and then it increases. To find this minimum, we take the derivative (recalling the chain rule) and get

$$D'(\alpha) = 2 - 4\frac{d\beta}{d\alpha}.$$

Remember that, from the Law of Refraction,

$$\frac{\sin\alpha}{\sin\beta} = k \text{ where } k = \frac{c_a}{c_w}.$$

If we differentiate this with respect to α we get

$$\cos\alpha = k\cos\beta \cdot \frac{d\beta}{d\alpha}.$$

Solving for $d\beta/d\alpha$ and substituting into the expression for $D'(\alpha)$ gives

$$D'(\alpha) = 2 - 4\frac{\cos\alpha}{k\cos\beta}.$$

Setting the derivative equal to zero we have,

$$D'(\alpha) = 2 - \frac{4}{k} \cdot \frac{\cos\alpha}{\cos\beta} = 0$$

$$\text{which implies} \quad \frac{k}{2} = \frac{\cos\alpha}{\cos\beta}.$$

We want the value of α which satisfies this equation, so we eliminate β. Squaring both sides gives

$$\frac{k^2}{4} = \frac{\cos^2 \alpha}{\cos^2 \beta} = \frac{\cos^2 \alpha}{1 - \sin^2 \beta}.$$

But $\sin \beta = \frac{1}{k} \sin \alpha$, so

$$\frac{k^2}{4} = \frac{\cos^2 \alpha}{1 - \frac{\sin^2 \alpha}{k^2}}$$

$$\frac{1}{4} = \frac{\cos^2 \alpha}{k^2 - \sin^2 \alpha} = \frac{\cos^2 \alpha}{k^2 - (1 - \cos^2 \alpha)}$$

$$k^2 - 1 + \cos^2 \alpha = 4 \cos^2 \alpha$$

$$\cos \alpha = \sqrt{\frac{k^2 - 1}{3}}.$$

Finally we have an expression for the cosine of the incidence angle with minimum deflection. Since raindrops are water, $k \approx 1.33$, so $\cos \alpha \approx 0.5063$ and $\alpha \approx 59.56°$. At this incidence angle, the deflection is $D(59.56) \approx 137.5°$. To establish that this is a minimum we can check the sign of the second derivative (see the exercises).

We have found the incidence angle, $\alpha \approx 59.58°$, that gives the minimum deflection. Since the derivative of the deflection function is zero at this special angle, we know that the change in deflection angle divided by the change in incidence angle is nearly zero near $\alpha \approx 59.58°$. In other words, many rays with incidence angle near 59.58° get deflected by about the same amount. Rays further away from this critical angle get spread out more. So if we are looking at the deflected light, then rays coming from the direction of minimum deflection should appear the brightest since they are spread out the least. This is where the rainbow appears. The ray whose incidence angle is $\alpha \approx 59.58°$ is called the *rainbow ray* and 42.5° (= 180 − 137.5) is called the *rainbow angle*. The rainbow angle is the angle from the horizontal at which an observer should see the rainbow, if the rays of sunlight are horizontal. Figure 4 shows how the rainbow angle is related to the sun, observer, and the raindrops.

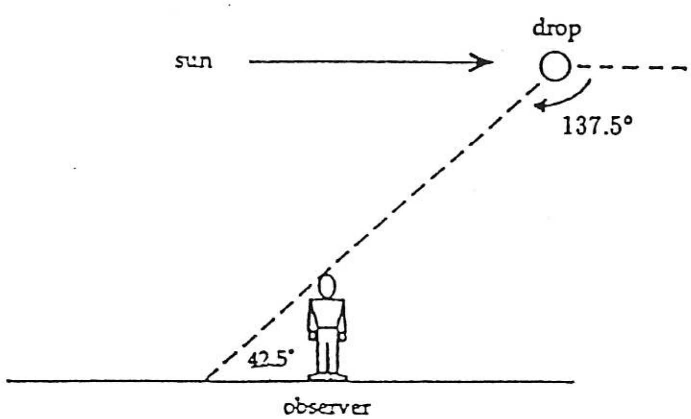

Figure 4.

Drops inclined 42.5° from an observer appear brighter than those with less inclination. For drops higher in the sky, the deflection angle would have to be less than 137.5° and since we discovered this angle is a minimum, no rays of the type we have been tracing come from drops higher in the sky. Any light coming from high in the sky must come from rays that have more than one (or none) internal reflections.

In the early part of the seventeenth century, Descartes carried out an analysis leading to discovery of the rainbow angle. Since the techniques of calculus were not available to him, he had to calculate the deflection of many different rays and even then did not have a nice expression for the incidence angle of minimum deflection.

Now that we know that light scattered by a drop is brighter at a certain angle of observation, any drop in the sky at the correct angle will show some brightness. Imagine the observer at the vertex of a cone with vertex angle equal to twice the rainbow angle. Cutting the cone with a plane perpendicular to its axis gives a circular cross-section and every raindrop on this circle forms the rainbow angle with the observer. Consequently, the observer should see a bright circular arc in the sky. This is the rainbow. Notice that the rainbow may be higher or lower in the sky depending on how high the sun is. To an observer on the ground, the rainbow is at most one half of a circle. However, to an observer flying in a plane, the rainbow may form an entire circle.

Exercises

4. Verify that we found a minimum deflection angle by checking the second derivative. (Hint: You can find the second derivative by first finding the second derivative of β with respect to α. The trigonometric formula for $\sin(\beta - \alpha)$ will be helpful.)
5. Sketch the function $D(\alpha)$ for α between 0 and 90 degrees.
6. If an observer sees the rainbow at an angle of 25 degrees from the horizontal, what is the sun's angle of inclination?

Colors

The geometry of light rays that we have considered so far accounts for a circular arc of brighter light in the sky. But where are the colors? Actually the answer is quite simple now. Light is really an electromagnetic wave and therefore we can talk about its frequency and wavelength. There is a wide spectrum of wavelengths, but our eyes are sensitive only to wavelengths in the range from about 7000 angstroms to about 4240 angstroms. Light with a wavelength of about 6470 to 7000 angstroms is perceived as red, and light in the 4000 to 4240 range is violet. Other colors fall between these two. Since the wave characteristic of these two colors are different, the refractive index of water varies depending on which color of light is passing through it. When red light with wavelength 6563 angstroms travels from air to water, the refractive index is about 1.3318. With violet light (4047 angstroms), the index increases to about 1.3435.

Sunlight is really a wide range of wavelengths. When it strikes a raindrop, wavelengths in the red range are refracted differently from those in the violet range. The other colors like blue and yellow fall between these two ranges and are refracted to various degrees between the two extremes. Consequently the light is actually spread into its constituent colors.

Now we need to repeat the calculation done to find the minimum angle of deflection. For red light, the minimum deflection is 137.7° and for violet light it is 139.4°. These values give rainbow angles of 42.3° and 40.6° respectively. In other words, when looking in the sky, the observer will see a circular arc of red light at a slightly higher inclination than the circular arc of violet light. The other wavelengths that we recognize as colors will form bows between these two. The order is red, orange, yellow, green, blue, indigo, and violet. (Taking the first letters gives a mnemonic: ROY G. BIV).

Newton was the first to make these careful calculations that explain the colors in the rainbow. By subtracting the rainbow angles for red and violet light it looks like the width of the bow is 1.7 degrees. Actually all these results assume that the rays from the sun are all parallel. To correct for the fact that the rays are not quite parallel, Newton allowed 0.5 degrees for the angular diameter of the sun and concluded that the rainbow width should be 2.2 degrees. This is in good agreement with actual observation although as we shall see later, the width of the bow does vary.

The Secondary Bow

Recall that the rainbow ray we traced was reflected once by the back of the raindrop. Other rays are reflected several times inside the drop. Each reflection reduces the intensity of the ray, but it is worth tracking at least those rays that have two internal reflections. To do this, look at Figure 5.

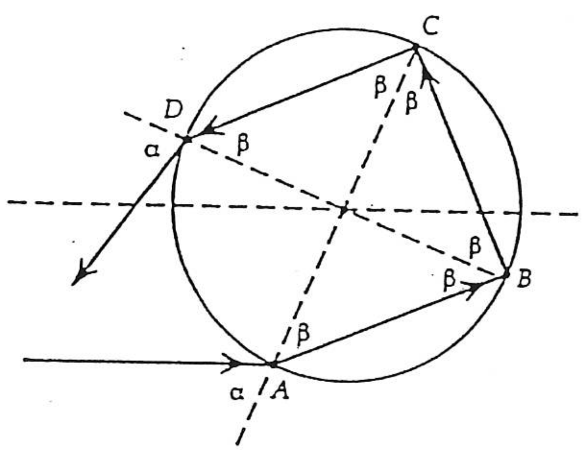

Figure 5.

This time we will follow rays incident on the bottom half of the drop since these rays are the ones that reach the observer. Keeping track of the deflections, we notice that the ray is rotated counter-clockwise at each of the points A, B, C, and D. The amount of rotation is similar to the analysis we did before, so this time we get

$$\begin{aligned}\text{Total deflection} &= (\alpha - \beta) + (180 - 2\beta) + (180 - 2\beta) + (\alpha - \beta)\\ &= 360 + 2\alpha - 6\beta.\end{aligned}$$

Since a 360 degree deflection means the ray continues in the same direction it started in, we can disregard the 360 and consider the deflection to be $2\alpha - 6\beta$. However, this is a counter-clockwise deflection and in order to compare it to the deflection for the rays with single internal reflections, we need to change this to a clockwise deflection. This is easily done by multiplying by -1. This gives us a new deflection function, D_2, for rays with two internal reflections:

$$D_2(\alpha) = 6\beta - 2\alpha.$$

Notice that $D_2(0) = 0$ and that D_2 begins to increase as α increases. In order to determine if this trend continues, we find any critical points by taking the derivative and setting it equal to zero. This time the critical point satisfies

$$\cos\alpha = \sqrt{\frac{k^2 - 1}{8}}.$$

With $k = 1.33$, we obtain the critical point $\alpha = 71.94°$, and $D_2(71.94) = 129.9°$. At this new critical point, D_2 is actually a maximum.

Hence for rays with two internal reflections, the maximum deflection angle is about 130°. In other words, raindrops that are inclined about 50°(i.e. 180° − 130°) from the observer will appear bright, although not as bright as those at 42°. This secondary arc of brightness is another bow which is dimmer that the primary bow and, unless conditions are right, is often too dim to see. Moreover, since D_2 is concave down, when we compare the maximum deflection for red light with that for violet light, we find that red light is deflected the most so the colors in the secondary bow appear in reverse order from those in the primary bow.

Notice also that the maximum of D_2 is about 130° while the minimum of D is about 138°. In other words, none of the rays with one or two internal reflections are deflected in the range 130 to 138 degrees. This means that the region between the primary and secondary bows is darker than the surrounding sky. It isn't totally black since light comes from rays that are reflected and refracted in many other ways. This darkened band is called Alexander's band after Alexander of Aphrodisias, a follower of Aristotle. Alexander deduced from Aristotle's theory of the rainbow that the region between the bows should be particularly bright. Since it wasn't, Alexander saw the need for a revised theory even though he couldn't supply one.

Exercises

7. Verify that the critical point for D_2 does occur at the point where $\cos\alpha = \sqrt{\frac{k^2 - 1}{8}}$.

8. Sketch the graph of D_2.

9. Determine the maximum deflection angle for red light and violet light.
10. Using the same procedure as above, find the deflection function D_n for rays that have n internal reflections. Find the critical point for this function. Theoretically, each of these classes of rays gives rise to another rainbow. They are rarely seen in the sky because they are so dim, but often one can see the first few bows in a laboratory set-up.

References

Boyer, Carl (1959), *The Rainbow from Myth to Mathematics*, Thomas Yoseloff Publishing, New York.

Greenler, Robert (1980), *Rainbows, Halos, and Glories*, Cambridge University Press, Cambridge.

Meyer-Arendt, Jurgen (1972), *Introduction to Classical and Modern Optics*, Prentice-Hall, New Jersey.

Nussenzveig, H. (1977), "The Theory of the Rainbow," *Scientific American*, April.

Tricker, R. (1970), *Introduction to Meteorological Optics*, American Elsevier, New York.

A longer version of this module appeared in *The UMAP Journal* 13: 149-174 (1992).