



ELSEVIER

Physics Reports 320 (1999) 27–36

PHYSICS REPORTS

www.elsevier.com/locate/physrep

From Alexander of Aphrodisias to Young and Airy

J.D. Jackson

University of California, Berkeley, CA, USA and Lawrence Berkeley National Laboratory, University of Berkeley, Berkeley, CA, 94720, USA

Abstract

A didactic discussion of the physics of rainbows is presented, with some emphasis on the history, especially the contributions of Thomas Young nearly 200 years ago. We begin with the simple geometrical optics of Descartes and Newton, including the reasons for Alexander's dark band between the main and secondary bows. We then show how dispersion produces the familiar colorful spectacle. Interference between waves emerging at the same angle, but traveling different optical paths within the water drops, accounts for the existence of distinct supernumerary rainbows under the right conditions (small drops, uniform in size). Young's and Airy's contributions are given their due. © 1999 Elsevier Science B.V. All rights reserved.

PACS: 01.30.Rr; 42.15.Dp; 42.25.Fx; 42.68.Ge

'This pedagogical piece on rainbows is dedicated to Lev B. Okun, colleague and friend, on his 70th birthday. On an extended visit to Berkeley in 1990, Lev saw on my office wall a picture of a double rainbow with at least three supernumerary bows visible inside the main bow. As part of my "lecture" on the photograph, I showed Lev a copy of these 1987 handwritten notes prepared for a class. He said, "Are these published somewhere?" My answer was no, but now they are, in augmented form. Lev is an amazing man, a physicist-mensch – a brilliant researcher, mentor, and warm human being. I have a vivid memory of a wonderful trip to Yosemite National Park with an allegedly ailing Lev. In the early morning hours, we found Lev outside our tent in Curry Village perched on a sloping rock doing vigorous calisthenics! Lev, may you have Many Happy Returns!'

The rainbow has fascinated since ancient times. Aristotle offered an explanation (not correct), as did clerics and scholars through the ages. Newton and Descartes established the elementary theory, according to what we now know as geometrical optics. But long before Newton and Descartes, as early as the 13th century, the puzzling occasional phenomenon of supernumerary rainbows was noted. These "aberrations" were inexplicable in terms of geometrical optics. It was not until the beginning of the 19th century that Thomas Young, promoting the wave theory of light against acolytes of Newton, offered the correct explanation of the supernumeraries as results of interference. Airy put the theory on a firm mathematical footing in 1836. A scholarly treatment of the

history of the attempts to understand the rainbow by Boyer [1] contains much of interest, including striking paintings and photographs with the rainbow as subject. A semi-popular account of the theory of rainbows is presented by Nussenzweig [2].

The discussion that follows traces the theory of the rainbow from the simple Cartesian–Newtonian description to the interference-diffraction-caustic treatment of Airy.

1. Geometrical optics, no dispersion

A light ray is incident on a water drop of radius a at impact parameter b , as shown in Fig. 1. The index of refraction of water at the wavelength of the sodium D lines ($\lambda = 5890, 5896 \text{ \AA}$) and at 20°C is closely $n = \frac{4}{3}$. The ray has an angle of incidence i whose sine is $\sin i = b/a \equiv x$. The angle of refraction r is given by Snell's law as $r = \sin^{-1}(x/n)$.

The scattering angle θ for the emerging light ray (defined here as the angle of emergence of the ray relative to the incident direction) can be computed by adding up the angular bends made by the ray:

The entering bend is $(i - r)$

Each internal reflection bend is $(\pi - 2r)$

The exiting bend is $(i - r)$

For m internal reflections, the scattering angle is thus

$$\theta_m = |2(i - r) + m(\pi - 2r)| \quad [\text{Modulo } 2\pi]. \quad (1)$$

The primary rainbow has $m = 1$, the secondary, $m = 2$, and so on. Fig. 2 shows the scattering angle as a function of $\sin i = b/a$ for $m = 1$ and $m = 2$. At the extremes, the angle is either 0 or π ,

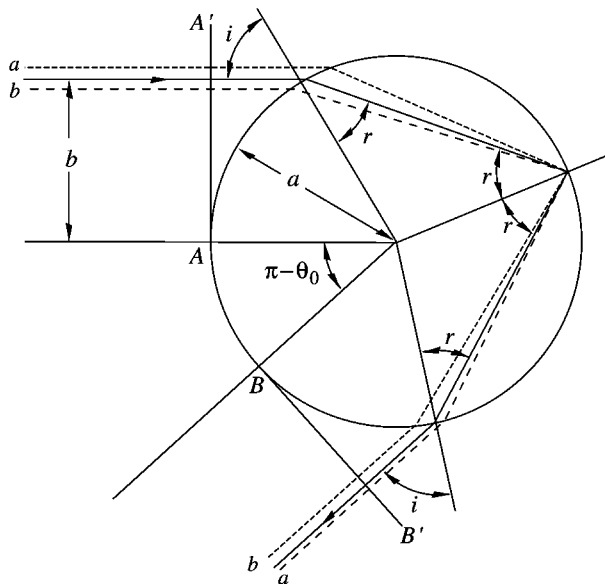


Fig. 1. Geometrical optics of a primary rainbow.

but for intermediate b/a values, the light is scattered at various angles. Note the gap between 129° and 138° . This is a region of negligible scattering (from higher orders) and appears as a dark space between the primary and secondary rainbows (known as Alexander’s dark space, after Alexander of Aphrodisias, a follower of Aristotle and head of the Lyceum in Athens around 200 AD).

The feature that causes the rainbow is the extremum in angle as a function of impact parameter. For the primary rainbow (upper curve in Fig. 2), the minimum angle is $\theta_0 = 138^\circ$ at $x_0 = 0.86066$ for $n = \frac{4}{3}$. Classically, the scattering cross section is

$$d\sigma/d\Omega = |b db/(\sin\theta d\theta)| \tag{2}$$

At the extremum, $db/d\theta$ is infinite, corresponding to a (classically) infinite cross section. Wave aspects prevent the infinity, of course, but it is indicative of a large cross section. The singular behavior is an example of a caustic.

To examine the vicinity of the extremum and see its dependence on the index of refraction, we make a Taylor series expansion around the minimum. For the primary bow ($m = 1$), we have

$$\theta = \pi + 2 \sin^{-1}x - 4 \sin^{-1}(x/n), \tag{3}$$

where $x = b/a$. The first two derivatives are

$$\frac{d\theta}{dx} = \frac{2}{\sqrt{1-x^2}} - \frac{4}{\sqrt{n^2-x^2}}, \tag{4}$$

$$\frac{d^2\theta}{dx^2} = \frac{2x}{(1-x^2)^{3/2}} - \frac{4x}{(n^2-x^2)^{3/2}}. \tag{5}$$

The extremum occurs for $d\theta/dx = 0$, i.e., $\sqrt{n^2-x_0^2} = 2\sqrt{1-x_0^2}$, or $x_0 = \sqrt{(4-n^2)/3}$ and $\sqrt{1-x_0^2} = \sqrt{(n^2-1)/3}$.

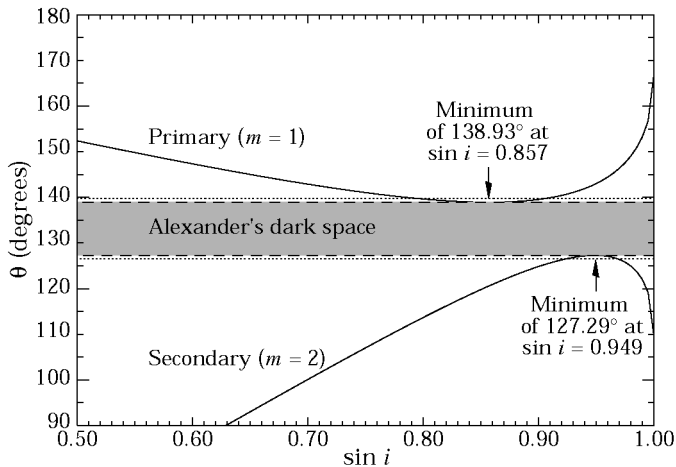


Fig. 2. Rainbow scattering angles according to geometrical optics for index of refraction $n = 1.34$. As indicated by the dotted lines, the dark band is somewhat wider in violet light, for which $n = 1.345$. Typical $m = 1$ and $m = 2$ rays are shown in Fig. 4(a).

The second derivative at $x = x_0$ is of interest:

$$\theta'' \equiv \left. \frac{d^2\theta}{dx^2} \right|_{x=x_0} = \frac{9}{2} \frac{\sqrt{4-n^2}}{(n^2-1)^{3/2}}. \quad (6)$$

For $n = \frac{4}{3}$, $\theta'' = 9.780$, and $\theta_0 = 137.97^\circ$.

For x near x_0 , we have $\theta \simeq \theta_0 + \theta''(x - x_0)^2/2$. In passing, we note that near $\theta = \theta_0$, the classical scattering cross section is

$$\frac{d\sigma}{d\Omega} \simeq a^2 \sqrt{\frac{2}{\theta''(\theta - \theta_0)}} \times \frac{x_0}{\sin \theta_0}. \quad (7)$$

As sketched in Fig. 4(b), scattering is concentrated at $\theta = \theta_0$, but also occurs for $\theta > \theta_0$. This is what causes the white appearance “inside” the primary bow (and “outside” the secondary bow).

2. Colors of the rainbow, dispersion

The beautiful colors of the rainbow are a consequence of the variation of the index of refraction of water with wavelength of the light. This dispersion, as it is called, is shown quantitatively in Fig. 3. If we arbitrarily define the visible range of wavelength to be from 400 nm (violet) to 700 nm (red), we find that the index of refraction differs by $\Delta n = 1.3 \times 10^{-2}$ from one end of the range to the other.

Now consider the effect of a change in n on θ :

$$\frac{d\theta}{dn} = -4 \frac{\partial}{\partial n} [\sin^{-1}(x/n)] = \frac{4x}{n\sqrt{n^2 - x^2}}. \quad (8)$$

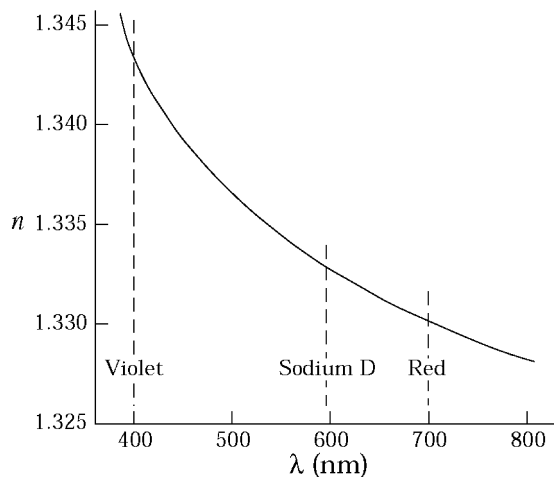


Fig. 3. Index of refraction of water as a function of wavelength. The visible light interval is between 400 and 700 nm.

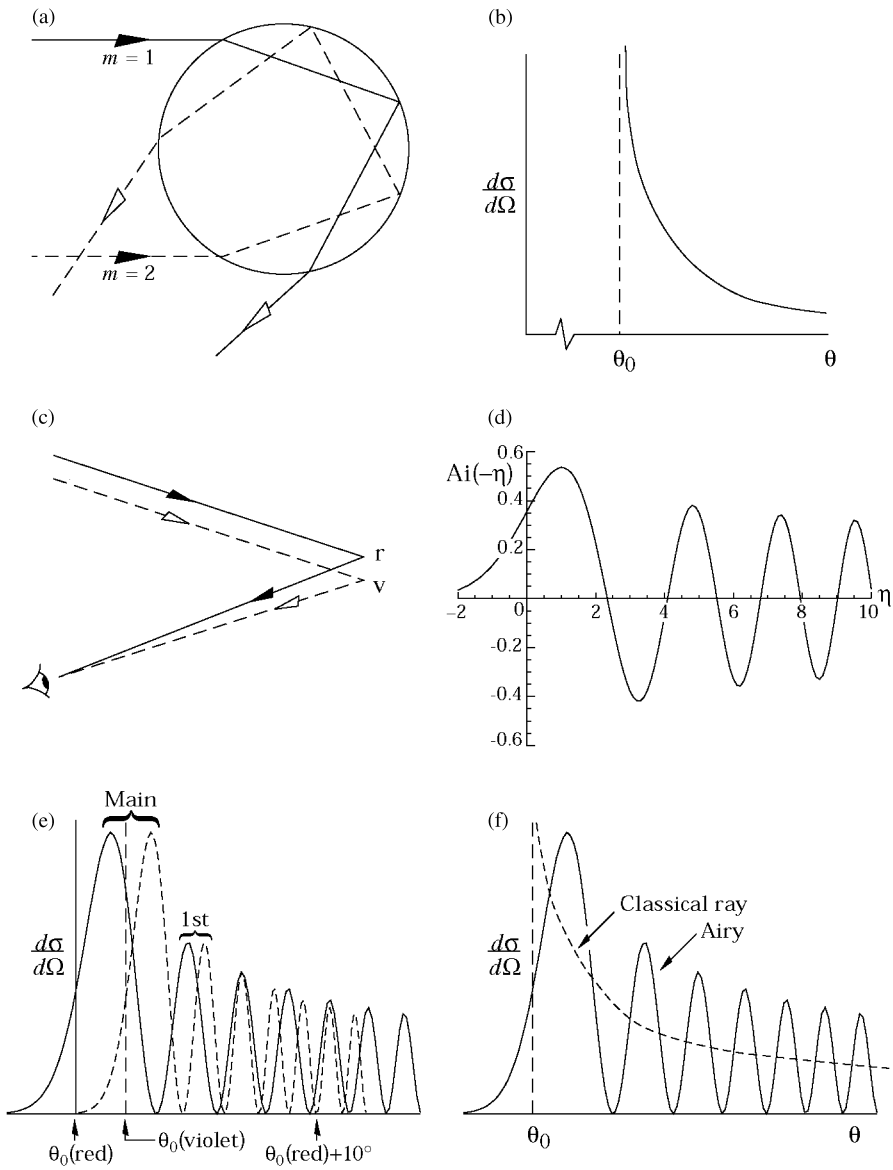


Fig. 4. Sketches to accompany the text.

At the rainbow angle,

$$\left(\frac{d\theta}{dn}\right)_{x_0} = \frac{2}{n} \sqrt{\frac{4-n^2}{n^2-1}} \tag{9}$$

For $n = 4/3$, $d\theta/dn|_{x_0} = 2.536$. With $\Delta n = 1.3 \times 10^{-2}$, we find $\Delta\theta_0 = 3.3 \times 10^{-2}$ radians = 1.89° . The colors of the rainbow are spread over about 2° out of the 42° away from the anti-solar point ($180^\circ - 138^\circ$). Since $dn/d\lambda < 0$, the red light emerges at a smaller angle than the violet.

The viewer thus sees the rainbow with the red at the outer side of the arc and the violet on the inner side, as indicated in Fig. 4(c). For the secondary bow, the order of the colors is opposite.

3. Consequences of the wave nature of light, supernumerary rainbows

For rays incident at impact parameters close to $b_0 = x_0 a$, the scattering angle is equal to θ_0 , correct to first order inclusive. In fact, because of the quadratic dependence of $\theta - \theta_0$ on $\Delta x = x - x_0$, two rays incident at impact parameters greater and less than b_0 by an amount $|\Delta x|$ will emerge with the *same* scattering angle. In the wave picture, as observed by Young [3] in 1803, these two waves emerging in the same direction can interfere. Whether the interference is constructive or destructive depends on the difference in optical path length of the two rays. This varies as a function of Δx and so provides the potential for interference effects in addition to dispersion in rainbows.

Referring to Fig. 1, we see as the solid line the critical ray, which emerges at $\theta = \theta_0$. On either side are shown neighboring rays with small $|\Delta x|$ that emerge at angles differing from θ_0 only in $O(\Delta x)^2$. The surfaces AA' and BB' are convenient ones for defining the optical path of a ray in the neighborhood of the critical ray. The optical path, or more appropriately, the phase accumulated along the ray, is given by

$$\phi(x) = 2ka(1 - \cos i + 2n \cos r), \quad (10)$$

where the $2(1 - \cos i)$ represents the sum of the distances from AA' to the drop's surface and similarly for the exit leg, while $4n \cos r$ is the length (times n) of the path interior to the drop. The free-space wave number is $k = \omega/c = 2\pi/\lambda$. In terms of $x = b/a$, the phase is

$$\phi(x) = 2ka[1 - \sqrt{1 - x^2} + 2\sqrt{n^2 - x^2}]. \quad (11)$$

We are interested in the behavior of $\phi(x)$ near $x - x_0$. Consider the derivative,

$$\frac{d\phi}{dx} = 2ka \left[\frac{x}{\sqrt{1 - x^2}} - \frac{2x}{\sqrt{n^2 - x^2}} \right]. \quad (12)$$

Comparison with $d\theta/dx$ in part(a) shows the relation,

$$d\phi/dx = kax d\theta/dx. \quad (13)$$

Writing $x = x_0 + \xi$, we can put this equation in the form,

$$d\phi/d\xi = ka[x_0 d\theta/d\xi + \xi d\theta/d\xi]. \quad (14)$$

Integration on both sides from 0 to ξ yields

$$\phi(\xi) - \phi_0 = ka \left[x_0(\theta - \theta_0) + \int_0^\xi \xi' \frac{d\theta}{d\xi'} d\xi' \right] \quad (15)$$

$$= ka \left[x_0(\theta - \theta_0) + \xi\theta - \int_0^\xi \theta(\xi') d\xi' \right]. \quad (16)$$

Inserting $\theta(\xi) = \theta_0 + \theta''\xi^2/2 + O(\xi^3)$ in the integral, we find

$$\phi(\xi) = \phi_0 + ka[x_0(\theta - \theta_0) + \theta''\xi^3/3 + O(\xi^4)]. \quad (17)$$

For two rays, a and b , as shown in Fig. 1, with equal and opposite small ξ values, the phase difference is

$$\delta = \phi(\xi) - \phi(-\xi) = 2ka\theta''\xi^3/3. \quad (18)$$

If we equate this phase difference to $2\pi N$ and express ξ in terms of $\theta - \theta_0$, we find the angles of constructive interference to be

$$\theta_N - \theta_0 \simeq [(\theta'')^{1/3}/2](3\pi N/ka)^{2/3}. \quad (19)$$

(Actually, a more correct procedure has $N + \frac{1}{4}$ replacing N . See Ref. [4, p. 243] and Section 3.21.)

The angles of constructive interference mark the positions of additional rainbows, called *supernumerary rainbows*. They lie at larger angles than θ_0 and so fall “inside” the main bow. Their colors are in the same order as in the primary bow. They are rarely seen because conditions must be optimized for them to appear unobscured or not washed out. The angle $(\theta_N - \theta_0)$ depends on the droplet size, varying as $(ka)^{-2/3}$. For large drops, the angle becomes very small and the supernumerary bows fall inside the various colors of the primary bow. Using $N = 5/4$, $\theta'' = 9.780$, and $(\theta - \theta_0)_{\min} = 3 \times 10^{-2}$ radians (corresponding to the spread caused by dispersion), we find $(ka)_{\max} \simeq 2.5 \times 10^3$. With k appropriate to the sodium D lines, we obtain $a_{\max} \simeq 0.28$ mm. Larger drops will cause the supernumeraries to be obscured by the effects of dispersion.

Variation in drop size, even if the drops are small enough, also causes the maxima of the supernumerary bows to be washed out in angle. Thus, one needs small drops, uniform in size, in order to see clearly the supernumeraries. All this was understood by Young [3].

For very small drop size, $a < 50 \mu\text{m}$, the whole pattern of primary peak ($N = \frac{1}{4}$) and supernumerary peaks for a given wavelength is so spread in angle that dispersion effects are unimportant. All the colors have broad primary peaks lying almost on top of each other in angle. The result is a “white rainbow” or “fog bow”.

4. Huygens’ construction for the rainbow, Airy integral

The scalar diffraction theory of Huygens, Young, Fresnel and Kirchhoff [5, Section 10.5] can be used to obtain an approximate description of the rainbow in wave theory, as was first done by George B. Airy (1836). Consider the line BB' in Fig. 1, where we have evaluated the expression for the phase $\phi(\xi)$. A wave along this line will have the form

$$\psi \propto \exp[ik_{\parallel}z + i\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp} + i\phi(\xi)], \quad (20)$$

where we choose our axes so that z is in the direction of scattering at θ_0 and \mathbf{r}_{\perp} is measured along BB' , with value ax_0 at the critical ray. If the wave is propagating in the direction θ , then

$$\mathbf{k}_{\perp} \cdot \mathbf{r}_{\perp} = -ka(\theta - \theta_0)x, \quad (21)$$

where the negative sign comes from the fact that \mathbf{k}_\perp and \mathbf{r}_\perp are antiparallel ($\theta > \theta_0$). Since z is constant on BB' , the relevant parts of the wave's overall phase are

$$\mathbf{k}_\perp \cdot \mathbf{r}_\perp + \phi(\xi) = ka[(\theta - \theta_0)(x_0 - x) + (\theta''/3)\xi^3] + \phi_0 \quad (22)$$

$$= ka[-\xi(\theta - \theta_0) + (\theta''/3)\xi^3] + \phi_0. \quad (23)$$

With the approximation, $\theta - \theta_0 = \theta''\xi^2/2$, we find

$$\mathbf{k}_\perp \cdot \mathbf{r}_\perp + \phi(\xi) - \phi_0 = -(ka/6)\theta''\xi^3 + O(\xi^4). \quad (24)$$

Along the line BB' , the wave amplitude in the neighborhood of $x = x_0$ or $\xi = 0$ has the form

$$\psi(\xi) = \exp(-ika\theta''\xi^3/6) \quad (25)$$

assuming the slowly varying amplitude function is a constant.

We can now use the simplest version of the Kirchoff integral for diffraction,

$$\psi_{\text{Scatt}} = \frac{k}{2\pi i} \int \frac{e^{ikR}}{R} \psi(\mathbf{x}') da'. \quad (26)$$

With $kR \simeq kr - \mathbf{k} \cdot \mathbf{x}'$ in the usual way, we find a scattering amplitude,

$$\psi_{\text{Scatt}} \propto \int_{-\infty}^{\infty} e^{ika(\theta - \theta_0)\xi} e^{-ika\theta''\xi^3/6} d\xi. \quad (27)$$

This can be put in the form of the *Airy integral* $\text{Ai}(-\eta)$, as defined by Abramowitz and Stegun [6, p. 447]:

$$\text{Ai}(-\eta) = \frac{1}{\pi} \int_0^{\infty} \cos(t^3/3 - \eta t) dt, \quad (28)$$

where $\eta = (2k^2 a^2 / \theta'')^{1/3} (\theta - \theta_0)$. This function is shown in Fig. 4(d). For positive η , $\text{Ai}(-\eta)$ oscillates with an amplitude that decreases as $\eta^{-1/4}$. For negative η , $\text{Ai}(-\eta)$ is exponential in character, falling rapidly to zero for $-\eta > 1$. The maxima and minima occur successively at

$$\eta = 1.0188 (1.1155), \quad 3.2482 (3.2616), \quad 4.8201 (4.8263), \quad 6.1633 (6.1671), \\ 7.3722 (7.3748).$$

The numbers in parentheses are values of $[(3\pi/2)(N + \frac{1}{4})]^{2/3}$ for $N = 0, 1, 2, \dots$, from our previous discussion of the angles for constructive interference. For larger N values, the agreement is excellent.

It is of interest to compare the angular positions of the supernumerary rainbows implicit in the tabulated η values with the examples quoted by Young [3]. Notorious for not giving details of his calculations, he only quotes answers. He states that for drops $\frac{1}{16}$ inches in diameter, the reds of the first and fourth supernumerary bows are approximately 2° and 4° inside the red of the primary (the first just clearing the violet of the primary). With $\theta'' = 9.912$ and $\lambda = 700$ nm for red light, we find $ka = 1.50 \times 10^3$ and $\eta = 1.34 (\theta - \theta_0)$, with the angles measured in degrees. With $\Delta\eta = 2.23$ and 6.35 for the first and fourth supernumeraries, we obtain $\Delta\theta = 1.7$ and 4.7 degrees, in rough agreement with Young. Incidentally, the fact that $\eta \neq 0$ for $N = 0$ (primary bow) explains the

long-standing puzzle that the angular positions of some rainbows were observed to vary appreciably away from the Cartesian–Newtonian angle θ_0 (evidently dependent on drop size).

The peak intensities of the supernumerary bows, relative to the primary bow, are 0.612, 0.504, 0.446, 0.408, ..., falling off only as $\eta^{-1/2}$ or $(N + \frac{1}{4})^{-1/3}$. Note that the $\eta^{-1/2}$ behavior is just what our classical cross section gave. The wave aspect rounds the corners and gives interference. The intensity pattern for red and violet light is sketched in Fig. 4(e) for $a = 64 \mu\text{m}$ ($ka \approx 10^3$ for violet light). The first supernumerary bow would be visible, but subsequent ones would not. Add some variation in drop size and everything except the primary bow will wash out.

An approximate cross section for a given ka can be written in terms of Airy’s integral by normalizing the *average* intensity at large η to the classical cross section. From Abramowitz and Stegun [6], one finds that for large η the leading term in an asymptotic series is

$$\text{Ai}(-\eta) \simeq (1/\sqrt{\pi})(1/\eta^{1/4})\sin(\frac{2}{3}\eta^{3/2} + \pi/4). \tag{29}$$

The average value of its square is $\langle |\text{Ai}(-\eta)|^2 \rangle = 1/(2\pi\sqrt{\eta})$. With the expression for η in terms of $\theta - \theta_0$, this becomes

$$\langle |\text{Ai}(-\eta)|^2 \rangle = \frac{1}{2\pi} \left(\frac{\theta''}{2}\right)^{2/3} \left(\frac{1}{ka}\right)^{1/3} \sqrt{\frac{2}{\theta''(\theta - \theta_0)}}. \tag{30}$$

Comparison with the classical cross section, near $\theta = \theta_0$,

$$\frac{d\sigma_{\text{cl}}}{d\Omega} \simeq a^2 \frac{x_0}{\sin \theta_0} \sqrt{\frac{2}{\theta''(\theta - \theta_0)}}, \tag{31}$$

leads to the cross section in the Airy approximation,

$$\frac{d\sigma}{d\Omega} \simeq 2\pi a^2 \frac{x_0}{\sin \theta_0} \left(\frac{2}{\theta''}\right)^{2/3} (ka)^{1/3} |\text{Ai}(-\eta)|^2. \tag{32}$$

(See Fig. 4(f).) For $n = \frac{4}{3}$, $\theta_0 = 137.97^\circ$, $\sin \theta_0 = 0.66952$, $x_0 = 0.86066$, and $\theta'' = 9.780$. Ignoring the loss of intensity from the refractions and reflection, the cross section for a given component of the rainbow of fixed ka is thus

$$\frac{d\sigma}{d\Omega} \simeq 2.80(ka)^{1/3} |\text{Ai}(-\eta)|^2 a^2. \tag{33}$$

At the peak of the rainbow, $d\sigma/d\Omega = 0.803(ka)^{1/3} a^2$. For $ka = 10^3$, this cross section is 32 times as great as an isotropic cross section, $d\sigma/d\Omega = a^2/4$.

5. Comment on polarizations and loss of intensity

After Young, but before Airy, David Brewster showed in 1812 that the scattered rainbow light was almost completely polarized, confirming earlier observations of Biot (of the Biot–Savart law in magnetism). The polarization comes about because the refracted and reflected intensities at each of the interfaces are different for different polarizations. The formulas of pp. 305–306 of Jackson [5]

can be used to show that at $\theta = \theta_0$, the ratios of scattered amplitude to incident amplitude are

$$\frac{E_{\text{Scatt}}}{E_{\text{inc}}} = \begin{cases} 8/27 & \text{for } E_{\perp} \text{ plane of incidence,} \\ 2\left(\frac{2n}{n^2+2}\right)^2\left(\frac{2-n^2}{2+n^2}\right) & \text{for } E_{\parallel} \text{ plane of incidence.} \end{cases} \quad (34)$$

For $n = \frac{4}{3}$, the intensity of perpendicular polarization is 8.78×10^{-2} of the incident, while the intensity of the parallel polarization *relative to the perpendicular* is 3.9×10^{-2} . The cross section quoted above must therefore be multiplied by approximately $\frac{1}{2} \times 1.039 \times 8.78 \times 10^{-2}$ for unpolarized light incident.

6. Note on notation

Van de Hulst [4] defines his Airy integral to be

$$f(z) = \int_0^{\infty} \cos\left[\frac{\pi}{2}(zt - t^3)\right] dt. \quad (35)$$

His z and our η are related by $z = (12/\pi^2)^{1/3}\eta$. His function $f(z)$ is $f(z) = (2\pi^2/3)^{1/3}\text{Ai}(-\eta)$. Note that other notations are used for the Airy integral. For example, see [7, Section 59].

Acknowledgements

Special thanks are owed to Gail Harper and Betty Armstrong, who T_EX'ed the equation-rich handwritten manuscript, and to Don Groom for generating postscript files from my hand-drawn figures.

References

- [1] C.B. Boyer, *The Rainbow*, Princeton University Press, Princeton, 1959, 1987.
- [2] H.M. Nussenzveig, *Scientific American*, April 1977, p. 116.
- [3] T. Young, Experiments and Calculations Relative to Physical Optics, Bakerian Lecture, November 24, 1803, publ. *Phil. Trans.* 1804, in: G. Peacock (Ed.), *Miscellaneous Works of Thomas Young*, Vol. I, John Murray, London, 1855, p. 179, esp. pp. 185–187.
- [4] H.C. Van de Hulst, *Light Scattering by Small Particles*, Wiley, New York, 1957, Chapter 13; also Dover reprint.
- [5] J.D. Jackson, *Classical Electrodynamics*, 3rd Edition, Wiley, New York, 1998.
- [6] M. Abramowitz, I. Stegun, eds., *Handbook of Mathematical Functions*, Dover, New York, 1972.
- [7] L.D. Landau, E.M. Lifshitz, *The Classical Theory of Fields*, Pergamon, Oxford, 1962.