

CHAPTER 1 REFLECTION AND REFRACTION

1.1 Introduction

This “book” is not intended to be a vast, definitive treatment of everything that is known about geometric optics. It covers, rather, the geometric optics of first-year students, whom it will either help or confuse yet further, though I hope the former. The part of geometric optics that often causes the most difficulty, particularly in getting the right answer for homework or examination problems, is the vexing matter of sign conventions in lens and mirror calculations. It seems that no matter how hard we try, we always get the sign wrong! This aspect will be dealt with in Chapter 2. The present chapter deals with simpler matters, namely reflection and refraction at a plane surface, except for a brief foray into the geometry of the rainbow. The rainbow, of course, involves refraction by a spherical drop. For the calculation of the radius of the bow, only Snell’s law is needed, but some knowledge of physical optics will be needed for a fuller understanding of some of the material in section 1.7, which is a little more demanding than the rest of the chapter.

1.2 Reflection at a Plane Surface

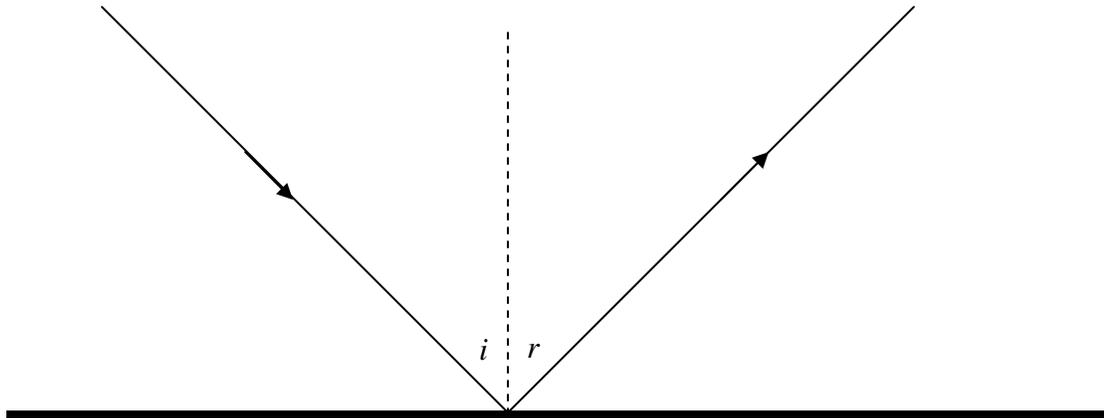


FIGURE I.1

The law of reflection of light is merely that the angle of reflection r is equal to the angle of incidence i . There is really very little that can be said about this, but I’ll try and say what little need be said.

i. It is customary to measure the angles of incidence and reflection from the normal to the reflecting surface rather than from the surface itself.

ii. Some curmudgeonly professors may ask for the lawS of reflection, and will give you only half marks if you neglect to add that the incident ray, the reflected ray and the normal are coplanar.

iii. A plane mirror forms a *virtual image* of a *real object*:

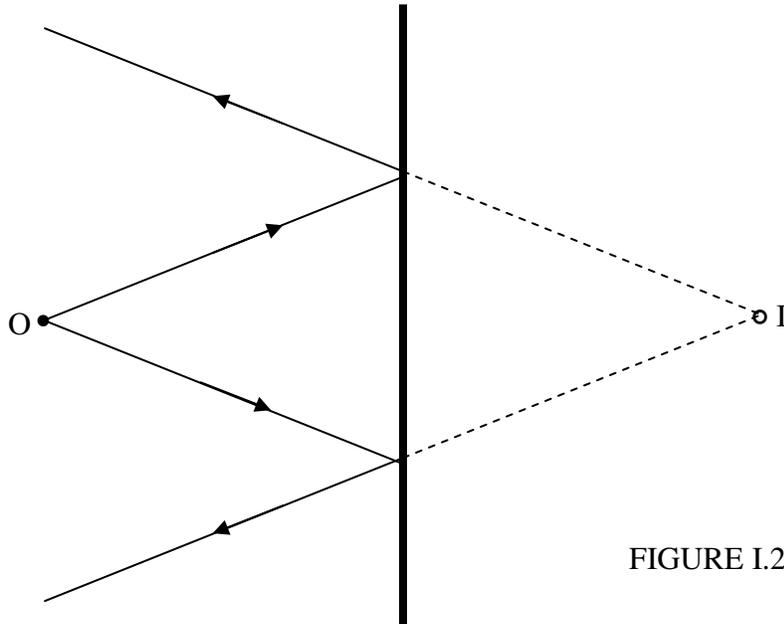


FIGURE I.2

or a *real image* of a *virtual object*:

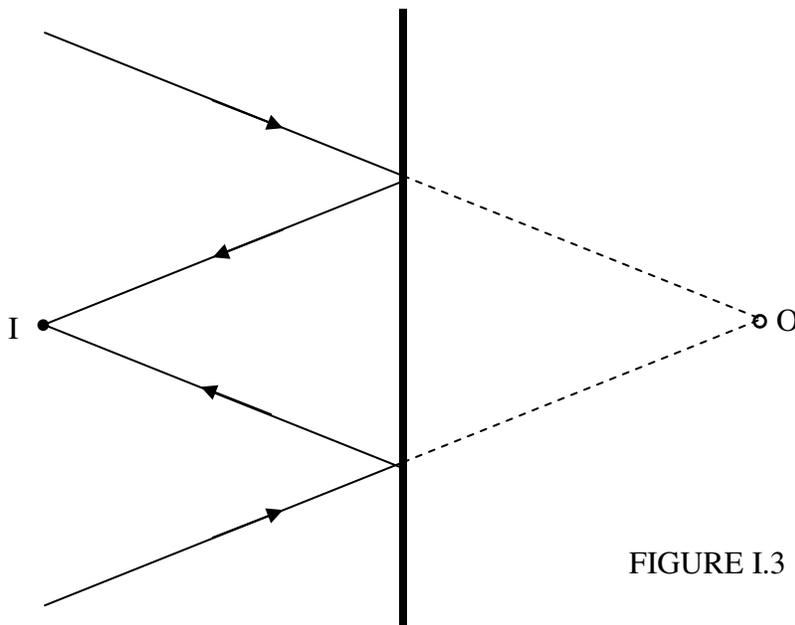


FIGURE I.3

iv. It is usually said that the image is as far behind the mirror as the object is in front of it. In the case of a virtual object (i.e. light converging on the mirror, presumably from some large lens somewhere to the left) you'd have to say that the image is as far *in front* of the mirror as the object is *behind it!*

v. If the mirror were to move at speed v away from a real object, the virtual image would move at speed $2v$. I'll leave you to think about what happens in the case of a virtual object.

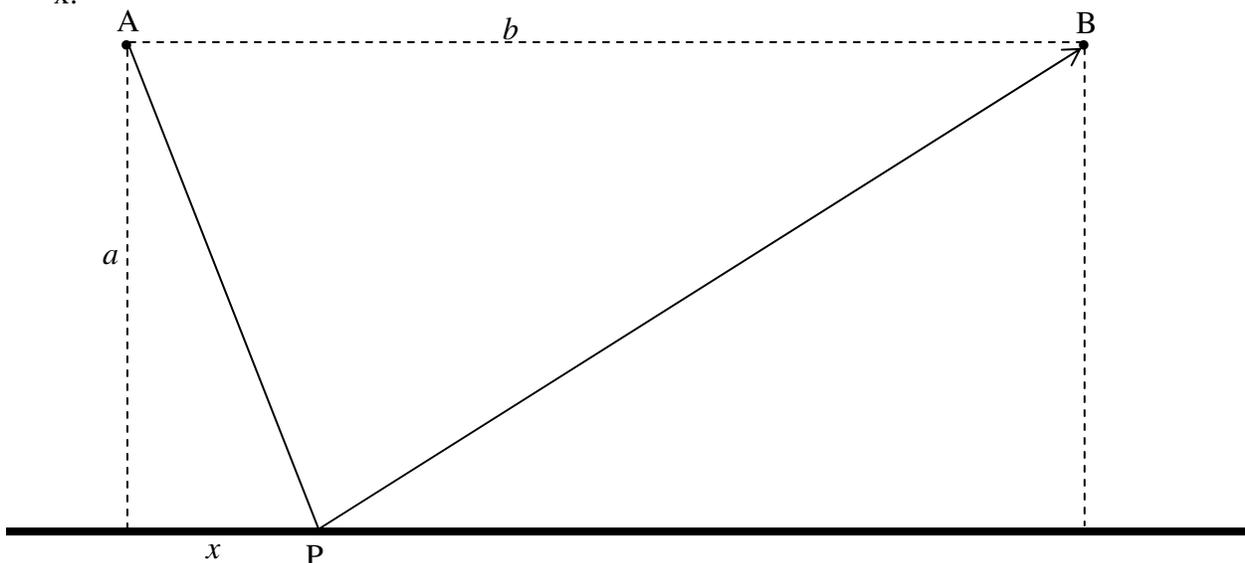
vi. If the mirror were to rotate through an angle θ (or were to rotate at an angular speed ω), the reflected ray would rotate through an angle 2θ (or at an angular speed 2ω).

vii. Only smooth, shiny surfaces reflect light as described above. Most surfaces, such as paper, have minute irregularities on them, which results in light being scattered in many directions. Various equations have been proposed to describe this sort of scattering. If the reflecting surface looks equally bright when viewed from all directions, the surface is said to be a perfectly diffusing Lambert's law surface. Reflection according to the $r = i$ law of reflection, with the incident ray, the reflected ray and the normal being coplanar, is called *specular* reflection (Latin: speculum, a mirror). Most surfaces are intermediate between specular reflectors and perfectly diffusing surfaces. This chapter deals exclusively with specular reflection.

viii. The image in a mirror is reversed from left to right, and from back to front, but is not reversed up and down. Discuss.

ix. If you haven't read *Through the Looking-glass and What Alice Found There*, you are missing something.

x.

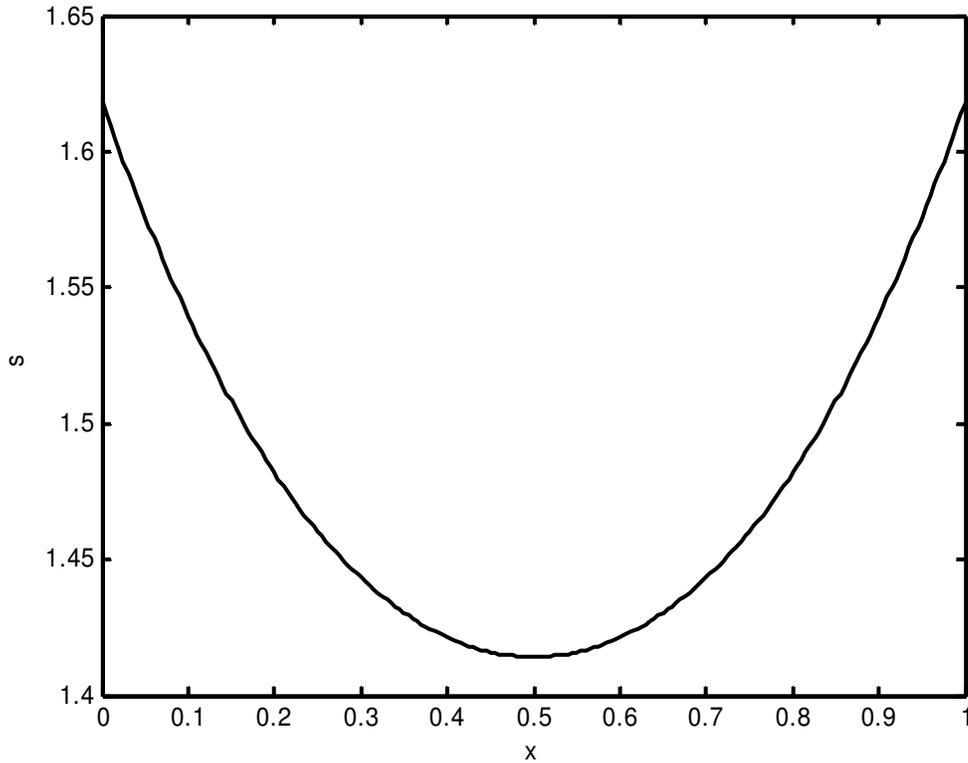


Light goes from A to B via reflection from a point P in a mirror.

The distance s travelled is given by

$$s = \sqrt{a^2 + x^2} + \sqrt{a^2 + (b-x)^2}. \quad 1.2.1$$

Here is the distance travelled as a function of the position of the point P:



The path that the light actually takes is the path such that the distance travelled is a minimum, which is such that P is horizontally halfway between A and B. You can see this from the graph, or by differentiating the above expression for s . This means that the angle of reflection is equal to the angle of incidence. You may regard this observation as a slightly interesting trivium, or as a fundamental principle of the deepest significance. Whichever you choose, you will come across lots of other examples of nature operating with Least Action. And you won't have to wait long. There's another one in the next section.

1.3 Refraction at a Plane Surface

I was taught Snell's Law of Refraction thus:

When a ray of light enters a denser medium it is refracted towards the normal in such a manner that the ratio of the sine of the angle of incidence to the sine of the angle of refraction is constant, this constant being called the *refractive index* n .

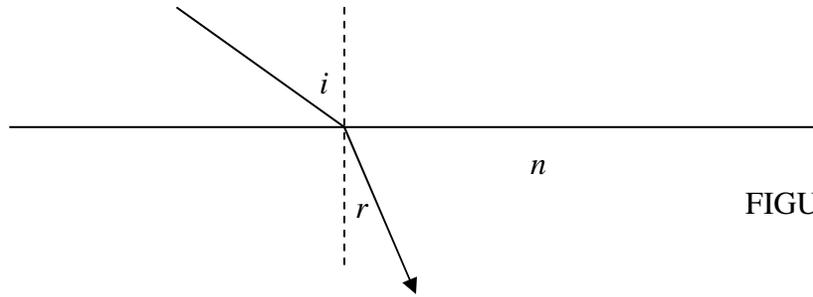


FIGURE I.4

This is all right as far as it goes, but we may be able to do better.

i. Remember the curmudgeonly professor who will give you only half marks unless you also say that the incident ray, the refracted ray and the normal are coplanar.

ii. The equation

$$\frac{\sin i}{\sin r} = n, \quad 1.3.1$$

where n is the refractive index of the medium, is all right as long as the light enters the medium from a vacuum. The refractive index of air is very little different from unity. Details on the refractive index of air may be found in my notes on Stellar Atmospheres (chapter 7, section 7.1) and Celestial Mechanics (subsection 11.3.3).

If light is moving from one medium to another, the law of refraction takes the form

$$n_1 \sin \theta_1 = n_2 \sin \theta_2. \quad 1.3.2$$

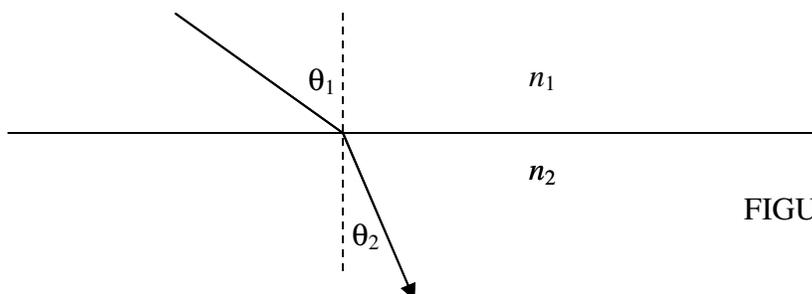


FIGURE I.5

iii. The statement of Snell's law as given above implies, if taken literally, that there is a one-to-one relation between refractive index and density. There must be a formula relating refractive index and density. If I tell you the density, you should be able to tell me the refractive index. And if I tell you the refractive index, you should be able to tell me the density. If you arrange substances in order of increasing density, this will also be their order of increasing refractive index.

This is not quite true, and, if you spend a little while looking up densities and refractive indices of substances in, for example, the *CRC Handbook of Physics and Chemistry*, you will find many examples of less dense substances having a higher refractive index than more dense substances. It is true in a general sense usually that denser substances have higher indices, but there is no one-to-one correspondence.

In fact light is bent towards the normal in a "denser" medium as a result of its slower speed in that medium, and indeed the speed v of light in a medium of refractive index n is given by

$$n = c/v, \quad 1.3.3$$

where c is the speed of light *in vacuo*. Now the speed of light in a medium is a function of the electrical permittivity ϵ and the magnetic permeability μ :

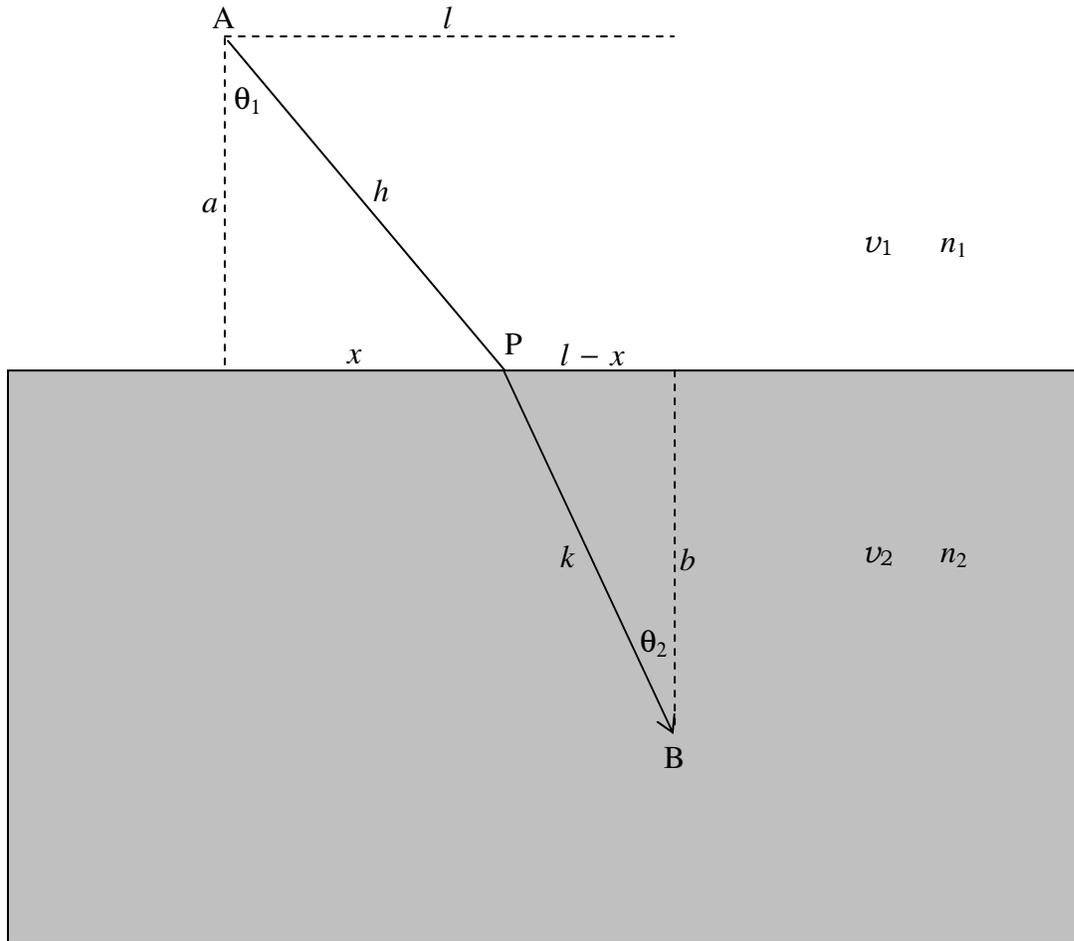
$$v = 1/\sqrt{\epsilon\mu}. \quad 1.3.4$$

The permeability of most nonferromagnetic media is very little different from that of a vacuum, so the refractive index of a medium is given approximately by

$$n \approx \sqrt{\frac{\epsilon}{\epsilon_0}}. \quad 1.3.5$$

Thus there is a much closer correlation between refractive index and relative permittivity (dielectric constant) than between refractive index and density. Note, however, that this is only an approximate relation. In the detailed theory there is a small dependence of the speed of light and hence refractive index on the frequency (hence wavelength) of the light. Thus the refractive index is greater for violet light than for red light (*violet* light is refracted more *violently*). The splitting up of white light into its constituent colours by refraction is called *dispersion*.

Here is a ray of light travelling from one medium to another:



It moves faster in the upper medium than in the lower medium.

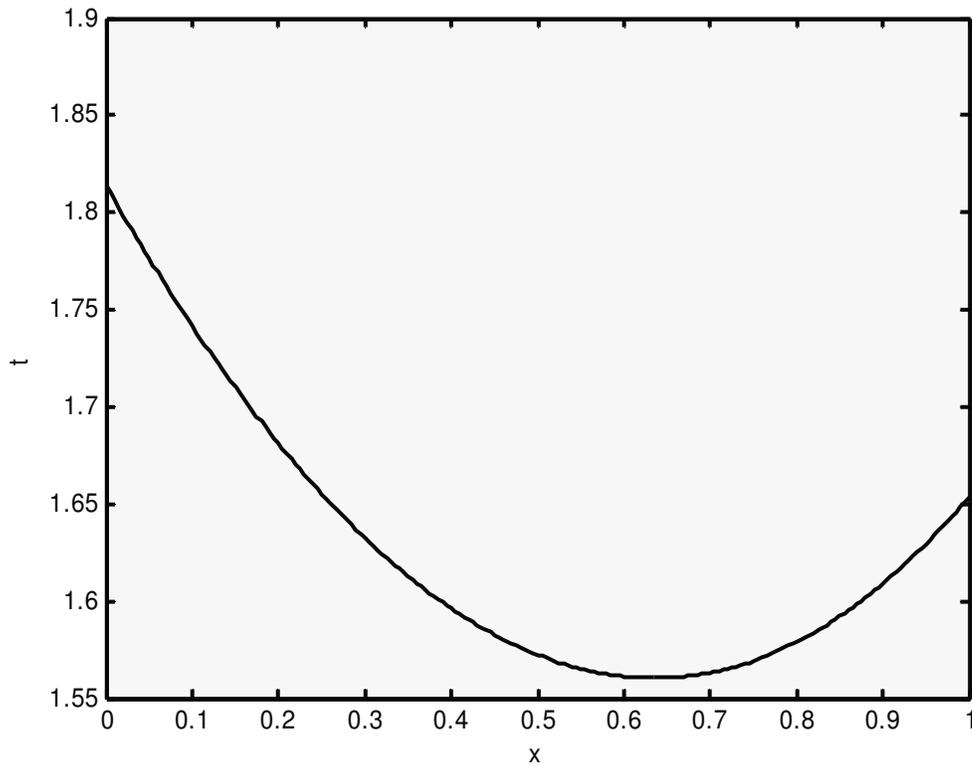
Time taken to get from A to B:

$$t = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (l-x)^2}}{v_2}. \quad 1.3.6$$

That is:

$$ct = n_1 \sqrt{a^2 + x^2} + n_2 \sqrt{b^2 + (l-x)^2}. \quad 1.3.7$$

Here is the time taken as a function of the position of P, calculated for $n_2/n_1 = 1.5$.



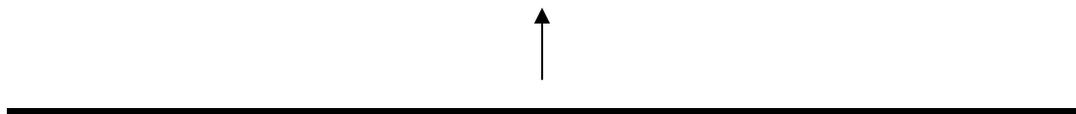
As you see, it goes through a minimum. You can find where it is by differentiating equation 1.3.7:

$$c \frac{dt}{dx} = \frac{n_1 x}{\sqrt{a^2 + x^2}} - \frac{n_2 (l - x)}{\sqrt{b^2 + (l - x)^2}} = n_1 \sin \theta_1 - n_2 \sin \theta_2, \quad 1.3.8$$

This is zero when $n_1 \sin \theta_1 - n_2 \sin \theta_2 = 0$. Thus Snell's law is such that the path actually taken is the path that takes the shortest time. Trivial, or profound?

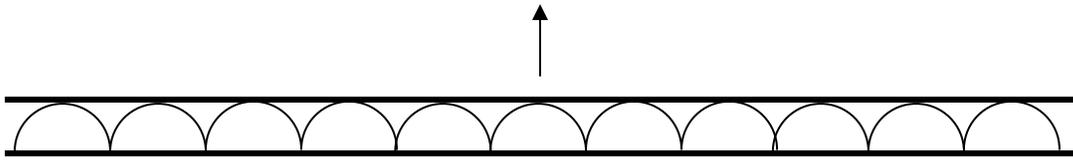
Huygens' Construction

Here is a wavefront moving upwards. "Light "rays" are normals to the wavefront.

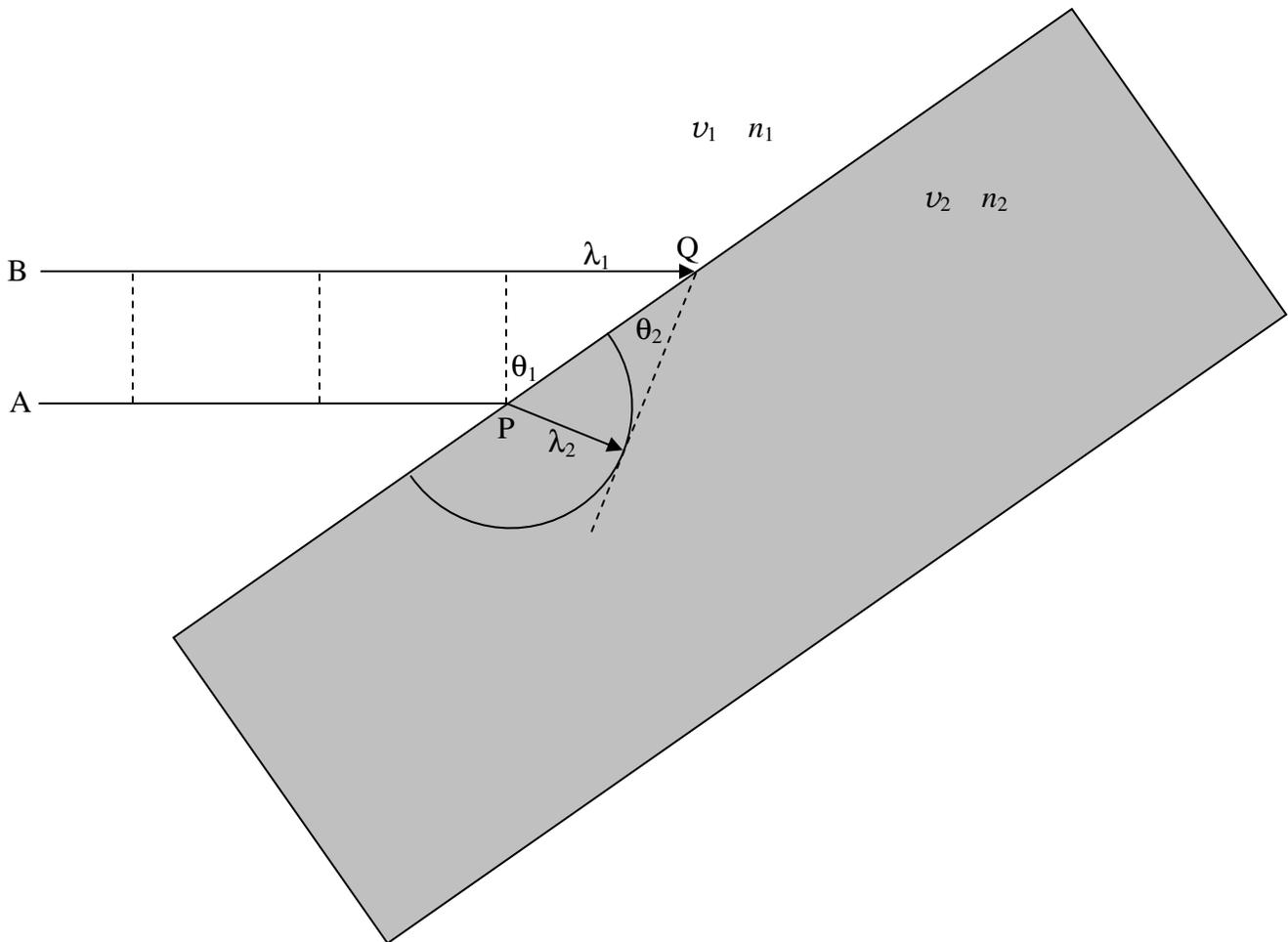


Huygens' construction is a way of prediction what will happen next. It says that you can imagine every point on the wavefront to be a source that generates a little wavelet. Then,

after a little time the wavelets look like this - and the new wavefront is the common tangent to all the wavelets.



This may sound trivial at first, although much has been written about it - i.e. whether it represents reality, or is merely a convenient construction. And, if real, what happens to the wavelets in the backwards direction? We'll not pursue that here, but we can use the Huygens construction as an interesting way to think about Snell's law.



A beam of light of wavelength λ_1 is approaching a glass block from the left at speed ν_1 . The dashed lines represent the wavefronts. Ray A reaches the block first, at P. A

wavelet is generated at P, moving with speed v_2 . The drawing is made for the instant when ray B reaches the point Q. The new wavefront is the tangent from Q to the little wavelet that started at P. The geometry will show that $\frac{\lambda_1}{\lambda_2} = \frac{v_1}{v_2} = \frac{n_2}{n_1}$ and therefore

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

Refraction through a glass block

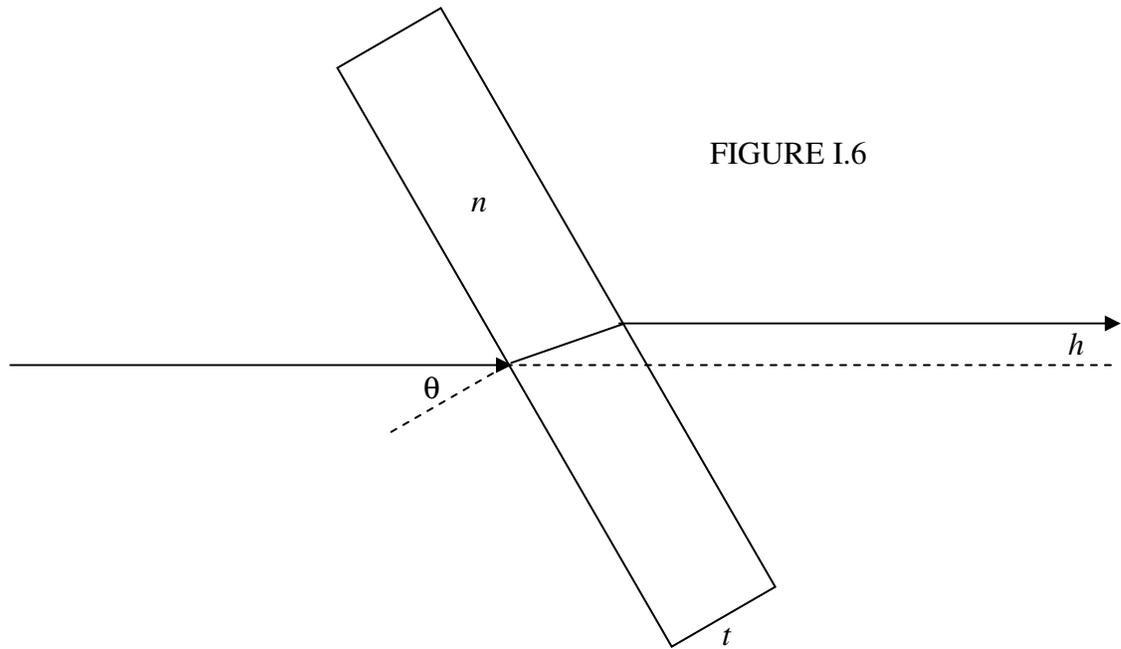


FIGURE I.6

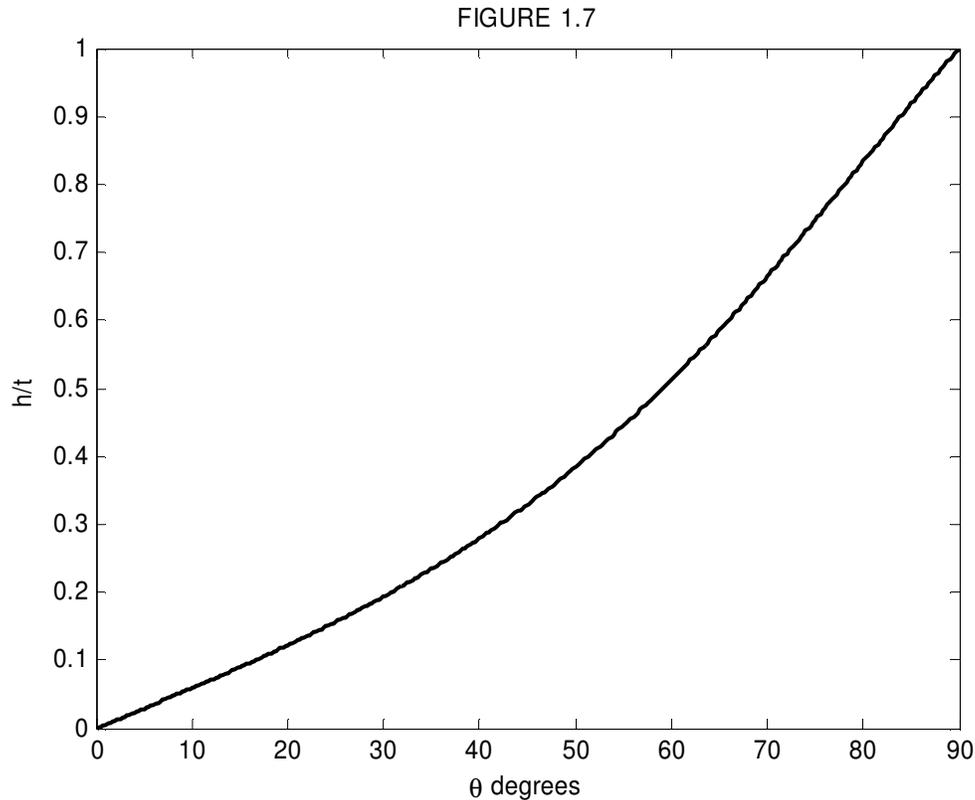
Figure I.6 shows a ray of light passing through a rectangular glass block of thickness t and refractive index n (taken to be 1.5 in the drawing). The normal to the surface of the block makes an angle θ with the incoming ray. It is a matter of simple geometry (do it!) to show that the *lateral displacement* h of the ray is given by

$$h = t(\sin \theta - \cos \theta \tan \phi), \quad 1.3.6$$

where ϕ is the angle of refraction, given by $\sin \theta = n \sin \phi$. In terms of θ , n and t , this is

$$h = t \sin \theta \left(1 - \frac{\cos \theta}{\sqrt{n^2 - \sin^2 \theta}} \right). \quad 1.3.7$$

Figure I.7 is drawn for $n = 1.5$.



One might imagine making use of this to measure the distance between two points close together. For example suppose that you have a photograph of some stars on an old photographic plate, and it includes a close pair of stars, and you want to measure the distance between the two star images. (Today the photograph would be on a CCD detector, and the distance between the two images would be recorded electronically, which is why I specify an old photographic plate.) You look at the photograph through a microscope and see one of the stars bisected by a crosshair in the microscope eyepiece. But you have a glass plate in front of the photograph, and you tilt the plate in order to displace the images so that the images move and the second star is now bisected by the crosshair. From the large angle through which you tilt the plate you can work out the tiny distance between the two images. You'll want to use monochromatic light, and you'd need to know the refractive index at that wavelength. How will you measure the angle through which the plate has turned? Well, you could shine a laser beam off it; the reflected light will move at twice the speed of the plate, and you could let it illuminate a sheet of graph paper several feet away. Thus the tiny distance between the images will correspond to a large distance on the graph paper. There may be one or two other practical details that you'd want to think about. For example, how thick would you want the glass plate to be? A thin microscope slide, maybe, or something much thicker than that? Would it be better to work at angles θ less than about 40° where the slope of figure I.7 is small, or at angles greater than 50° where the slope is larger?

You might also wish to move a laser beam sidewise through a small and controlled amount. You could put a glass block on a turntable which could be rotated through a tiny measurable angle and thus move the laser beam laterally and accurately through a very tiny amount.

Let's continue with the glass block. Show that, if the glass block were to be rotated counterclockwise with angular speed ω , the laser beam would move upwards at a speed

$$\dot{h} = t \left[\cos \theta - \frac{n^2 - 2n^2 \sin^2 \theta + \sin^4 \theta}{(n^2 - \sin^2 \theta)^{3/2}} \right] \omega. \quad 1.3.8$$

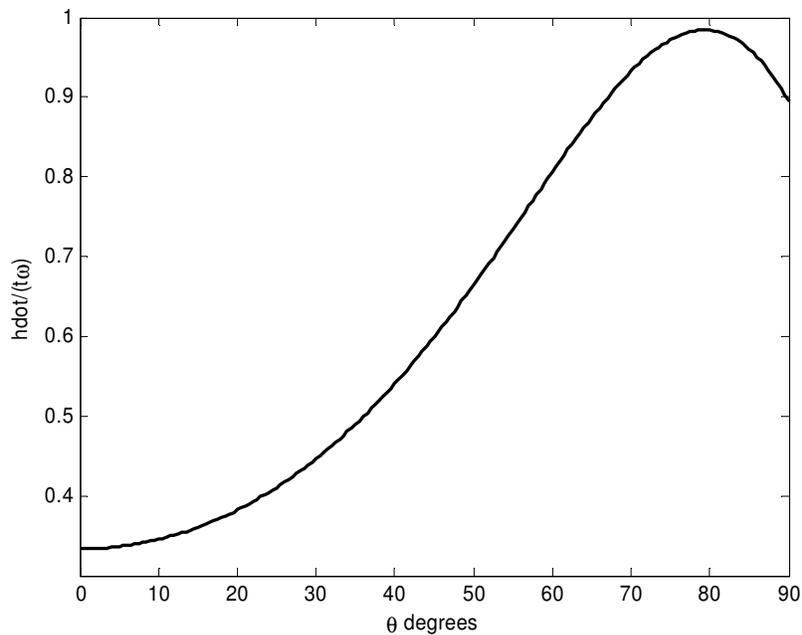
If we write $\sin^2 \theta$ as $1 - \cos^2 \theta$, we'll be able to express this entirely in terms of $\cos \theta$. And if, for illustrative purposes, I take $n = 1.5$, the equation becomes

$$\frac{\dot{h}}{t\omega} = c + \frac{1.25 - 2.5c^2 - c^4}{(1.25 + c^2)^{3/2}}, \quad 1.3.9$$

where $c = \cos \theta$. For ease in computation, I'll set $C = c^2$. Equation 1.3.9 then becomes

$$\frac{\dot{h}}{t\omega} = c + \frac{1.25 - C(2.5 + C)}{(1.25 + C)^{3/2}}. \quad 1.4.0$$

This is easy to compute, and the result is shown in the graph below.



We see, perhaps to our surprise, that \dot{h} goes through a maximum at about $\theta = 79^\circ$. To obtain equation 1.3.9, we had to differentiate equation 1.3.7. Now, to find out where \dot{h} goes through a maximum, we are going to have to differentiate again, although mercifully we can differentiate equation 1.4.0 with respect to C rather than to θ . If we do this, and then set the derivative to zero, we find, after some simplification,

$$\frac{1}{2C^{1/2}} = \frac{5 + 1.25C + 0.5C^2}{(1.25 + C)^{5/2}} \quad 1.4.1$$

If we square this and collect powers of C , we arrive at a quartic equation in C :

$$625 - 17980C - 6240C^2 - 2176C^3 + 256C^4 = 0.$$

where $C = c^2 = \cos^2 \theta$.

The solution to this equation is $C = 0.034\ 346\ 549\ 0$, corresponding to

$$\theta = \underline{79.319\ 731\ 1}$$

1.4 *Real and Apparent Depth*

When we look down into a pool of water from above, the pool looks less deep than it really is. Figure I.6 shows the formation of a virtual image of a point on the bottom of the pool by refraction at the surface.

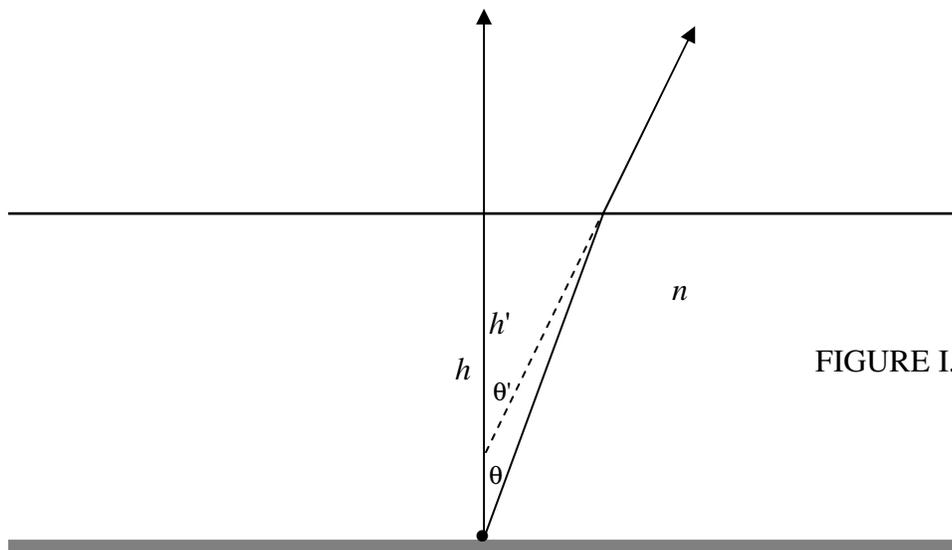


FIGURE I.6

The diameter of the pupil of the human eye is in the range 4 to 7 mm, so, when we are looking down into a pool (or indeed looking at anything that is not very close to our eyes), the angles involved are small. Thus in figure I.6 you are asked to imagine that all the angles are small; actually to draw them small would make for a very cramped drawing. Since angles are small, I can approximate Snell's law by

$$n \approx \frac{\tan \theta'}{\tan \theta} \quad 1.4.1$$

and hence

$$\frac{\text{real depth}}{\text{apparent depth}} = \frac{h}{h'} = \frac{\tan \theta'}{\tan \theta} = n. \quad 1.4.2$$

For water, n is about $\frac{4}{3}$, so that the apparent depth is about $\frac{3}{4}$ of the real depth.

Exercise. An astronomer places a photographic film, or a CCD, at the primary focus of a telescope. He then decides to insert a glass filter, of refractive index n and thickness t , in front of the film (or CCD). In which direction should he move the film or CCD, and by how much, so that the image remains in focus?

Now if Snell's law really were given by equation 1.4.1, all refracted rays from the object would, when produced backwards, appear to diverge from a single point, namely the virtual image. But Snell's law is really $n = \frac{\sin \theta'}{\sin \theta}$, so what happens if we do not make the small angle approximation?

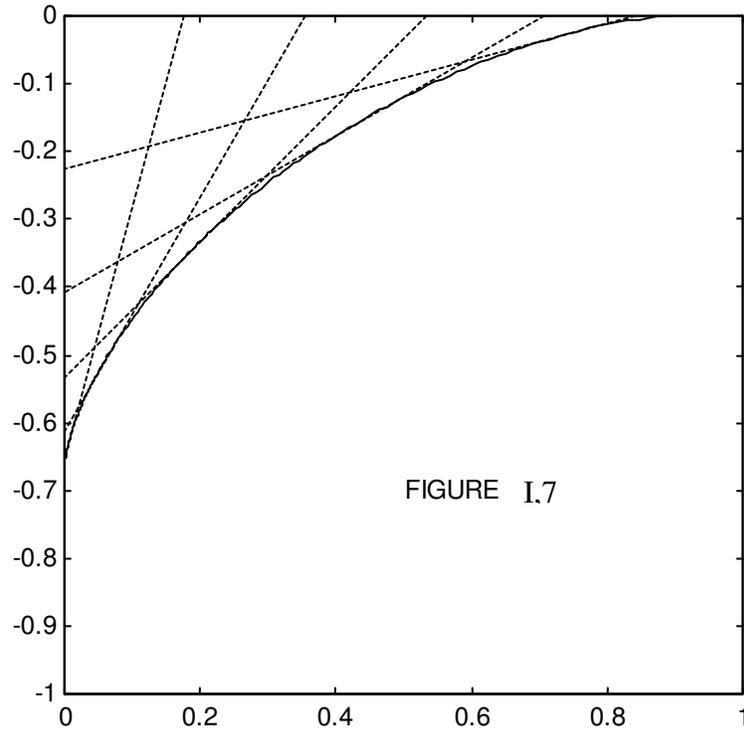
We have $\frac{h}{h'} = \frac{\tan \theta'}{\tan \theta}$ and, if we apply the trigonometric identity $\tan \theta = \frac{\sin \theta}{\sqrt{1 - \sin^2 \theta}}$

and apply Snell's law, we find that

$$\frac{h}{h'} = \frac{n \cos \theta}{\sqrt{1 - n^2 \sin^2 \theta}}. \quad 1.4.3$$

Exercise. Show that, to first order in θ this becomes $h/h' = n$.

Equation 1.4.3 shows h' as a function of θ – and that the refracted rays, when projected backwards, do not all appear to come from a single point. In other words, a point object does not result in a point image. Figure I.7 shows (for $n = 1.5$ – i.e. glass rather than water) the backward projections of the refracted rays for $\theta' = 15, 30, 45, 60$ and 75 degrees, together with their envelope or “caustic curve”. The “object” is at the bottom left corner of the frame, and the surface is the upper side of the frame.



Exercise (for the mathematically comfortable). Show that the parametric equations for the caustic curve are

$$x - y \tan \theta' - h \tan \theta = 0 \quad 1.4.4$$

and

$$ny \sec^3 \theta' + h \sec^3 \theta = 0. \quad 1.4.5$$

Here, $y = 0$ is taken to be the refracting surface, and θ and θ' are related by Snell's law.

Thus refraction at a plane interface produces an *aberration* in the sense that light from a point object does not diverge from a point image. This type of aberration is somewhat similar to the type of aberration produced by reflection from a spherical mirror, and to that extent the aberration could be referred to as "spherical aberration". If a point at the bottom of a pond is viewed at an angle to the surface, rather than perpendicular to it, a further aberration called "astigmatism" is produced. This will be discussed in Chapter 4.

1.5 Reflection and Refraction

We have described reflection and refraction, but of course when a ray of light encounters an interface between two transparent media, a portion of it is reflected and a portion is refracted, and it is natural to ask, even during an early introduction to the subject, just what fraction is reflected and what fraction is refracted. The answer to this is quite complicated, for it depends not only on the angle of incidence and on the two refractive indices, but also on the initial state of polarization of the incident light; it takes us quite far into electromagnetic theory and is beyond the scope of this chapter, which is intended to deal largely with just the *geometry* of reflection and refraction. However, since it is a natural question to ask, I can give explicit formulas for the fractions that are reflected and refracted in the case where the incident light is unpolarized.

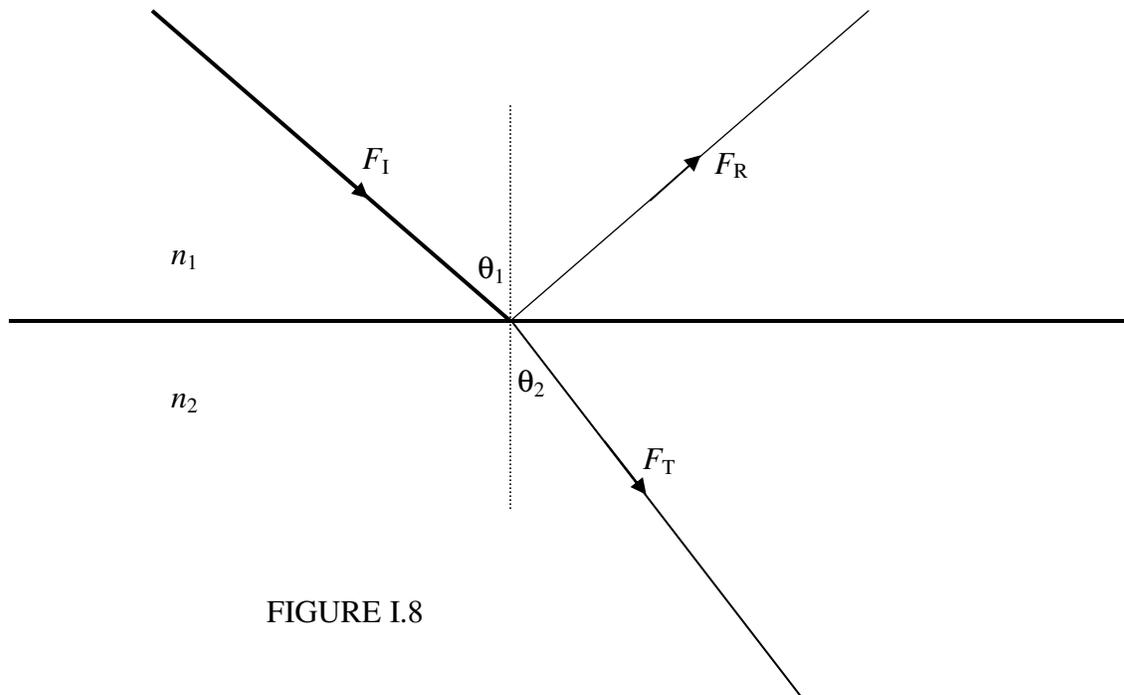


FIGURE I.8

Figure I.8 shows an incident ray of energy flux density (W m^{-2} normal to the direction of propagation) F_1 arriving at an interface between media of indices n_1 and n_2 . It is subsequently divided into a reflected ray of flux density F_R and a transmitted ray of flux density F_T . The fractions transmitted and reflected (t and r) are

$$t = \frac{F_T}{F_1} = 2n_1n_2 \cos \theta_1 \cos \theta_2 \left(\frac{1}{(n_1 \cos \theta_1 + n_2 \cos \theta_2)^2} + \frac{1}{(n_1 \cos \theta_2 + n_2 \cos \theta_1)^2} \right) \quad 1.5.1$$

and

$$r = \frac{F_R}{F_I} = \frac{1}{2} \left[\left(\frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \right)^2 + \left(\frac{n_1 \cos \theta_2 - n_2 \cos \theta_1}{n_1 \cos \theta_2 + n_2 \cos \theta_1} \right)^2 \right]. \quad 1.5.2$$

Here the angles and indices are related through Snell's law, equation 1.3.2. If you have the energy, show that the sum of these is 1.

Both the transmitted and the reflected rays are partially plane polarized. If the angle of incidence and the refractive index are such that the transmitted and reflected rays are perpendicular to each other, the reflected ray is completely plane polarized – but such details need not trouble us in this chapter.

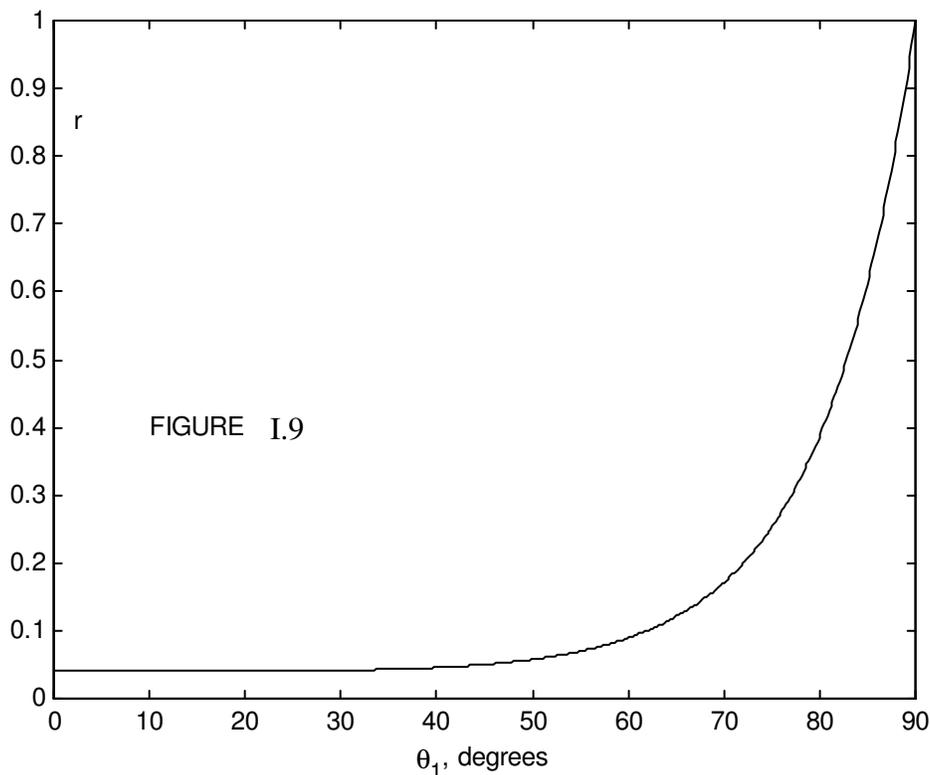


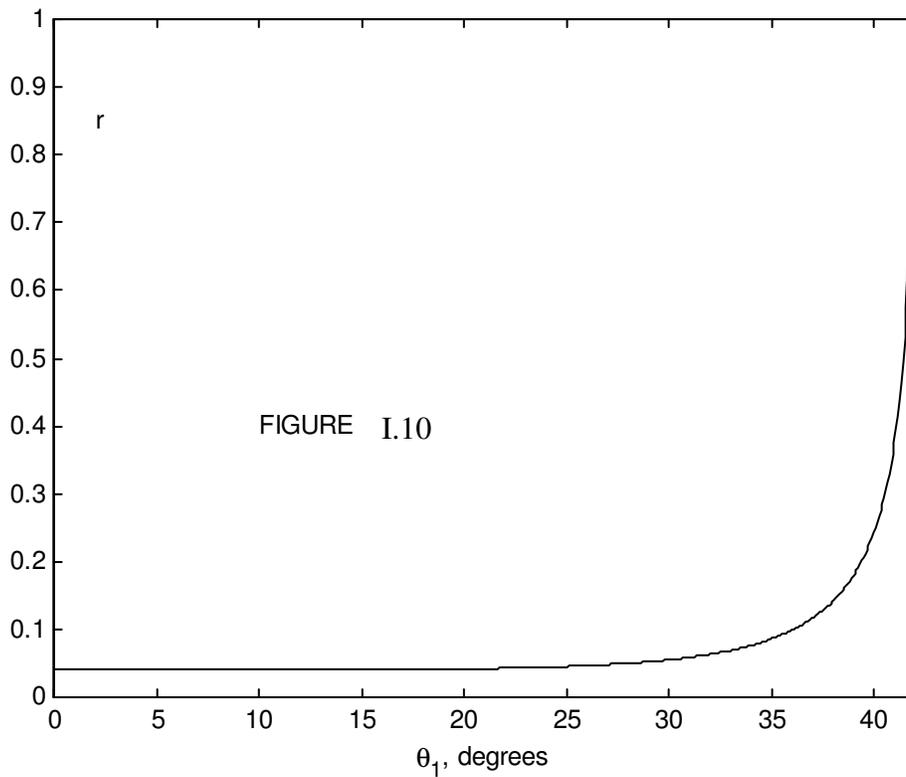
Figure I.9 shows the reflection coefficient as a function of angle of incidence for unpolarized incident light with $n_1 = 1.0$ and $n_2 = 1.5$ (e.g. glass). Since $n_2 > n_1$, we have *external reflection*. We see that for angles of incidence less than about 45 degrees, very little of the light is reflected, but after this the reflection coefficient increases rapidly with angle of incidence, approaching unity as $\theta_1 \rightarrow 90^\circ$ (grazing incidence).

If $n_1 = 1.5$ and $n_2 = 1.0$, we have *internal reflection*, and the reflection coefficient for this case is shown in figure I.10. For internal angles of incidence less than about 35° , little

light is reflected, the rest being transmitted. After this, the reflection coefficient increases rapidly, until the internal angle of incidence θ_1 approaches a *critical angle* C , given by

$$\sin C = \frac{n_2}{n_1}, \quad 1.5.3$$

This corresponds to an angle of emergence of 90° . For angles of incidence greater than this, the light is *totally internally reflected*. For glass of refractive index 1.5, the critical angle is $41^\circ.2$, so that light is totally internally reflected inside a 45° prism such as is used in binoculars.



1.6 Refraction by a Prism

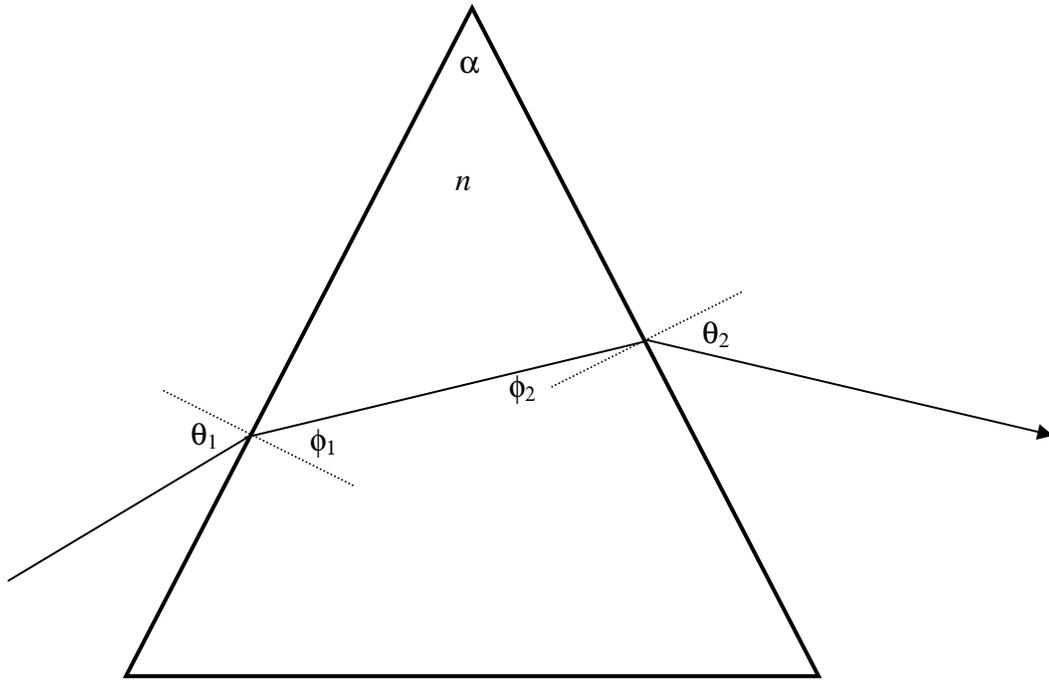


FIGURE I.11

Figure I.11 shows an isosceles prism of angle α , and a ray of light passing through it. I have drawn just one ray of a single colour. For white light, the colours will be *dispersed*, the violet light being deviated by the prism more than the red light. We'll choose a wavelength such that the refractive index of the prism is n . The *deviation* D of the light from its original direction is $\theta_1 - \phi_1 + \theta_2 - \phi_2$. I want to imagine, now, if we keep the incident ray fixed and rotate the prism, how does the deviation vary with angle of incidence θ_1 ? By geometry, $\phi_2 = \alpha - \phi_1$, so that the deviation is

$$D = \theta_1 + \theta_2 - \alpha. \quad 1.6.1$$

Apply Snell's law at each of the two refracting surfaces:

$$\frac{\sin \theta_1}{\sin \phi_1} = n \quad \text{and} \quad \frac{\sin \theta_2}{\sin(\alpha - \phi_1)} = n, \quad 1.6.2a,b$$

and eliminate ϕ_1 :

$$\sin \theta_2 = \sin \alpha \sqrt{n^2 - \sin^2 \theta_1} - \cos \alpha \sin \theta_1. \quad 1.6.3.$$

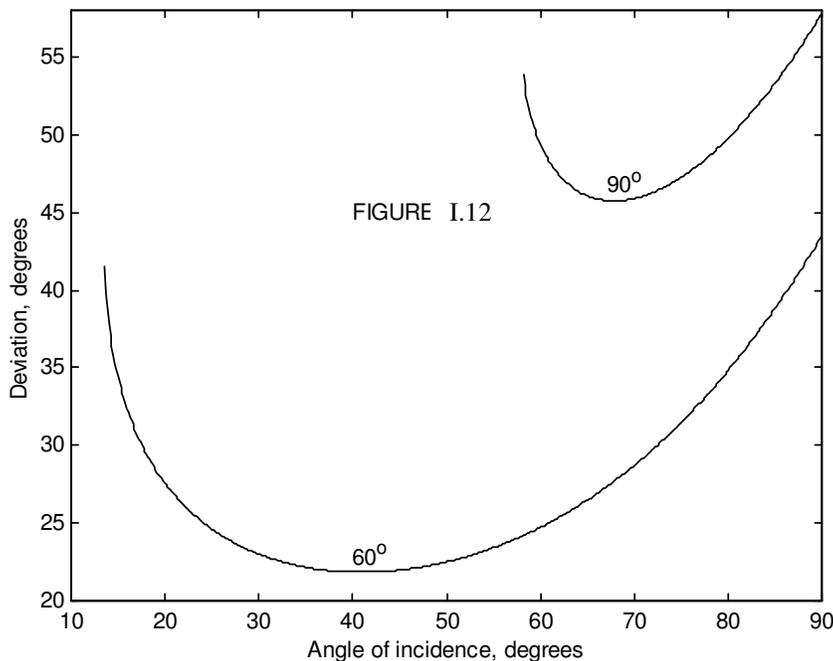
Equations 1.6.1 and 1.6.3 enable us to calculate the deviation as a function of the angle of incidence θ_1 . The deviation is least when the light traverses the prism symmetrically, with $\theta_1 = \theta_2$, the light inside the prism then being parallel to the base. Putting $\theta_1 = \theta_2$ in equation shows that minimum deviation occurs for an angle of incidence given by

$$\sin \theta_1 = \frac{n \sin \alpha}{\sqrt{2(1 + \cos \alpha)}} = n \sin \frac{1}{2} \alpha . \quad 1.6.4$$

The angle of minimum deviation D_{\min} is $2\theta_1 - \alpha$, where θ_1 is given by equation 1.6.4, and this leads to the following relation between the refractive index and the angle of minimum deviation:

$$n = \frac{\sin \frac{1}{2}(D_{\min} + \alpha)}{\sin \frac{1}{2} \alpha} . \quad 1.6.5$$

Of particular interest are prisms with $\alpha = 60^\circ$ and $\alpha = 90^\circ$. I have drawn, in figure I.12 the deviation versus angle of incidence for 60- and 90-degree prisms, using (for reasons I shall explain) $n = 1.31$, which is approximately the refractive index of ice. For the 60° ice prism, the angle of minimum deviation is $21^\circ.8$, and for the 90° ice prism it is $45^\circ.7$.



The geometry of refraction by a regular hexagonal prism is similar to refraction by an equilateral (60°) triangular prism (figure I.13):

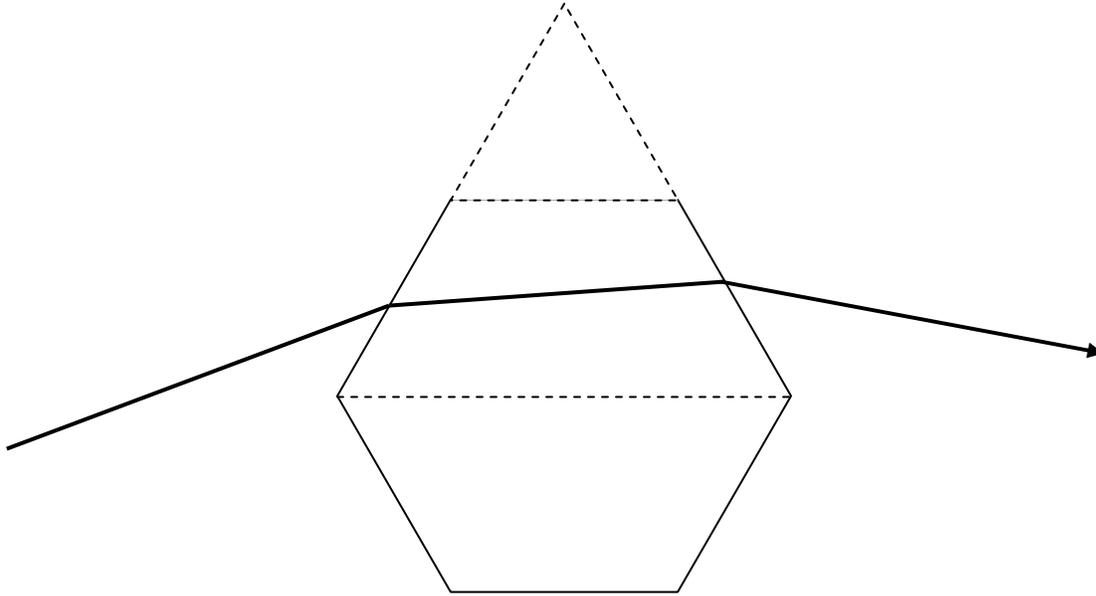


FIGURE I.13

When hexagonal ice crystals are present in the atmosphere, sunlight is scattered in all directions, according to the angles of incidence on the various ice crystals (which may or may not be oriented randomly). However, the rate of change of the deviation with angle of incidence is least near minimum deviation; consequently much more light is deviated by $21^\circ.8$ than through other angles. Consequently we see a halo of radius about 22° around the Sun.

Seen sideways on, a hexagonal crystal is rectangular, and consequently refraction is as if through a 90° prism (figure I.14):

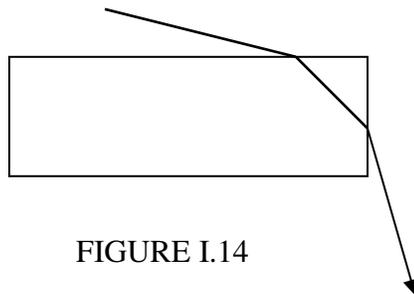


FIGURE I.14

Again, the rate of change of deviation with angle of incidence is least near minimum deviation, and consequently we may see another halo, of radius about 46° . For both haloes, the violet is deviated more than the red, and therefore both haloes are tinged violet on the outside and red on the inside.

1.7 The Rainbow

I do not know the exact shape of a raindrop, but I doubt very much if it is drop-shaped! Most raindrops will be more or less spherical, especially small drops, because of surface tension. If large, falling drops are distorted from an exact spherical shape, I imagine that they are more likely to be flattened to a sort of horizontal pancake shape rather than drop-shaped. Regardless, in the analysis in this section, I shall assume drops are spherical, as I am sure small drops will be.

We wish to follow a light ray as it enters a spherical drop, is internally reflected, and finally emerges. See figure I.15. We shall refer to the distance b as the *impact parameter*.

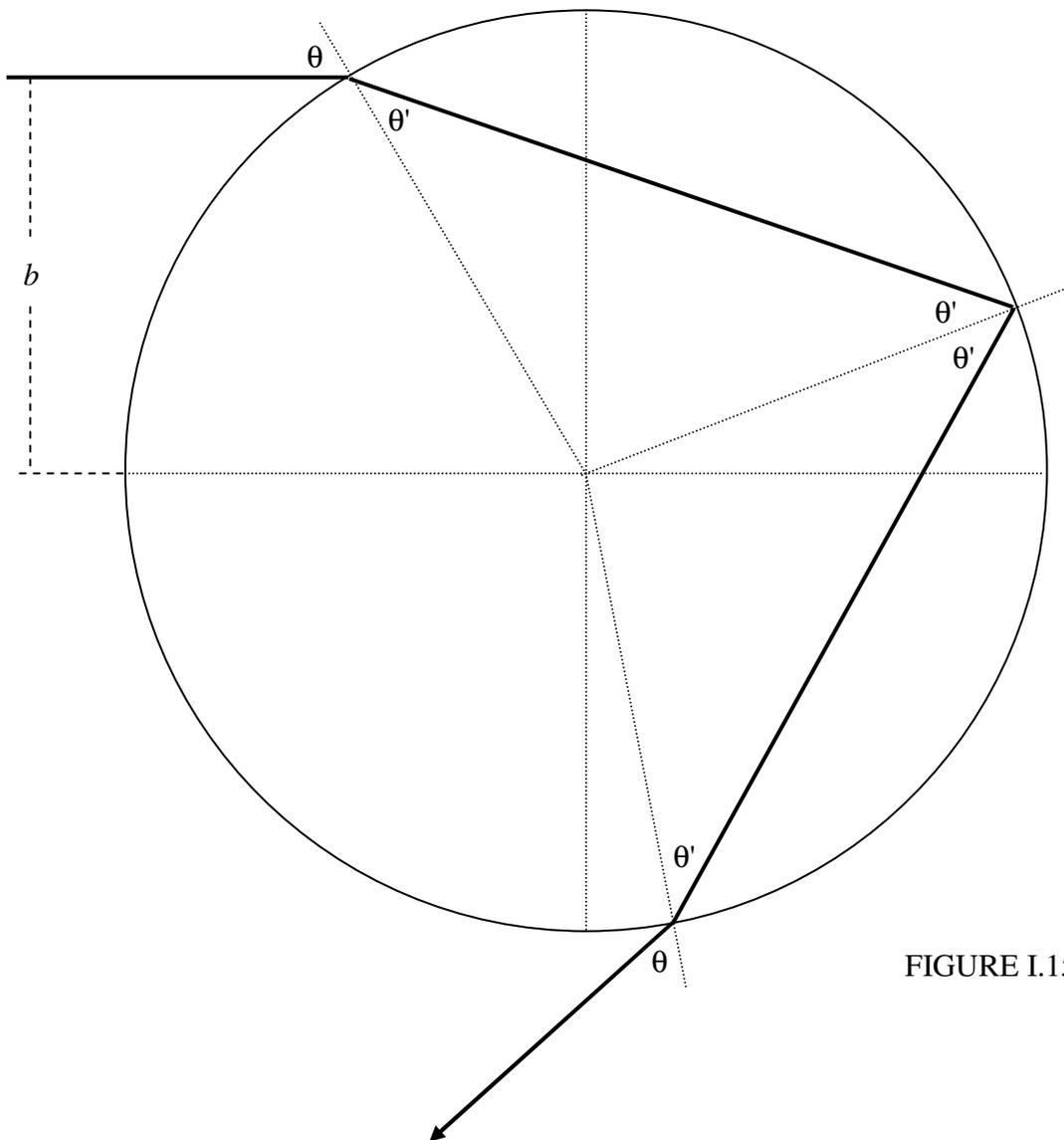


FIGURE I.15

We see a ray of light headed for the drop, which I take to have unit radius, at impact parameter b . The deviation of the direction of the emergent ray from the direction of the incident ray is

$$D = \theta - \theta' + \pi - 2\theta' + \theta - \theta' = \pi + 2\theta - 4\theta'. \quad 1.7.1$$

However, we shall be more interested in the angle $r = \pi - D$. A ray of light that has been deviated by D will approach the observer from a direction that makes an angle r from the centre of the bow, which is at the anti-solar point (figure I.16):

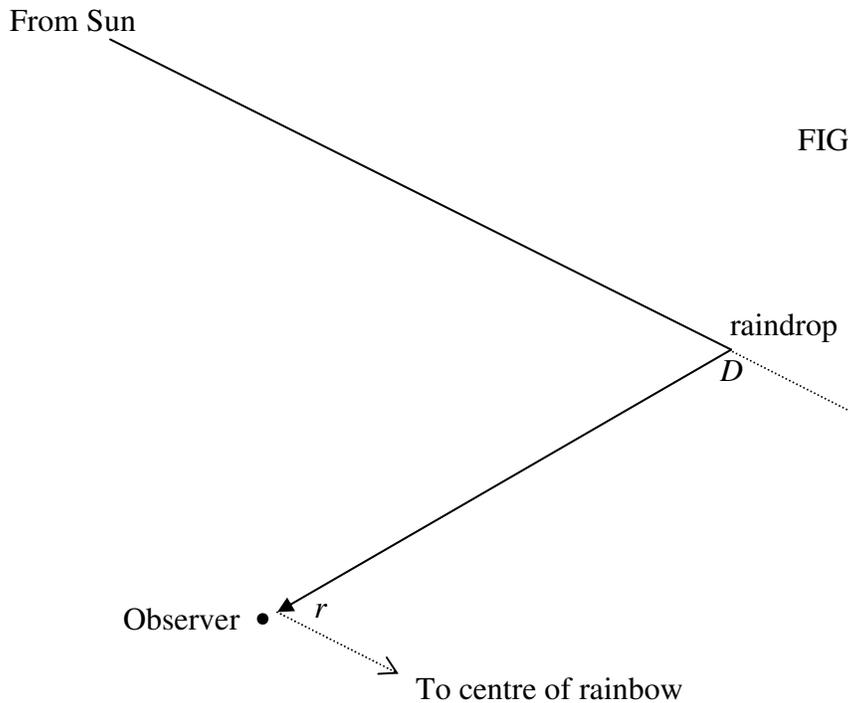


FIGURE I.16

We would like to find the deviation D as a function of impact parameter. The angles of incidence and refraction are related to the impact parameter as follows:

$$\sin \theta = b, \quad 1.7.2$$

$$\cos \theta = \sqrt{1 - b^2}, \quad 1.7.3$$

$$\sin \theta' = b/n, \quad 1.7.4$$

and
$$\cos \theta' = \sqrt{1 - b^2/n^2}. \quad 1.7.5$$

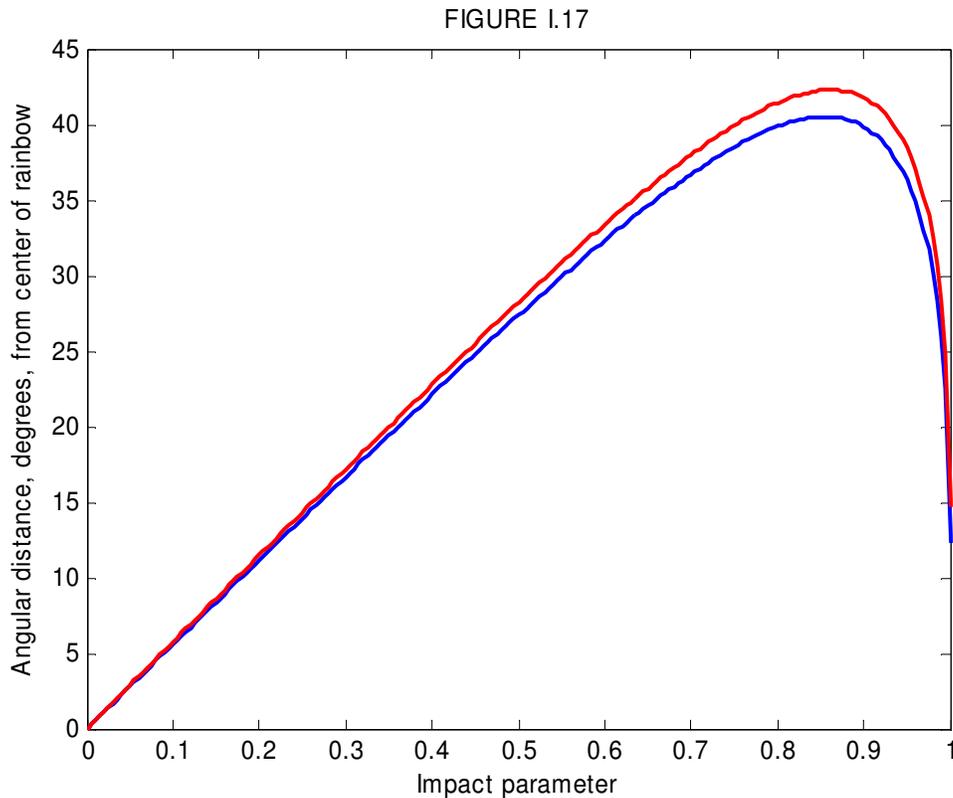
These, together with equation 1.7.1, give us the deviation as a function of impact parameter. The deviation goes through a minimum – and r goes through a maximum. The deviation for a light ray of impact parameter b is

$$D = \pi + 2\sin^{-1}b - 4\sin^{-1}(b/n). \quad 1.7.6$$

The angular distance r from the centre of the bow is $r = \pi - D$, so that

$$r = 4\sin^{-1}(b/n) - 2\sin^{-1}b. \quad 1.7.7$$

This is shown in figure I.17 for $n = 1.3439$ (blue - $\lambda = 400$ nm) and $n = 1.3316$ (red - $\lambda = 650$ nm).



Differentiation gives the maximum value, R , of r - i.e. the radius of the bow – or the minimum deviation D_{\min} . We obtain for the radius of the bow

$$R = 4\sin^{-1}\sqrt{\frac{4-n^2}{3n^2}} - 2\sin^{-1}\sqrt{\frac{4-n^2}{3}}. \quad 1.7.8$$

For $n = 1.3439$ (blue) this is $40^\circ 31'$ and for $n = 1.3316$ (red) this is $42^\circ 17'$. Thus the blue is on the inside of the bow, and red on the outside.

For grazing incidence (impact parameter = 1), the deviation is $2\pi - 4\sin^{-1}(1/n)$, or $167^\circ 40'$ for blue or $165^\circ 18'$ for red. This corresponds to a distance from the centre of the bow $r = 4\sin^{-1}(1/n) - \pi$, which is $12^\circ 20'$ for blue and $14^\circ 42'$ for red. It will be seen from figure I.17 that for radii less than R (i.e. inside the rainbow) but greater than $12^\circ 20'$ for blue and $14^\circ 42'$ for red there are *two* impact parameters that result in the same deviation, i.e. in the same position inside the bow. The paths of two rays with the same deviation are shown in figure I.18. One ray is drawn as a full line, the other as a dashed line. They start with different impact parameters, and take different paths through the drop, but finish in the same direction. The drawing is done for a deviation of 145° , or 35° from the bow centre. The two impact parameters are 0.969 and 0.636. When these two rays are recombined by being brought to a focus on the retina of the eye, they have satisfied all the conditions for interference, and the result will be brightness or darkness according as to whether the path difference is an even or an odd number of half wavelengths.

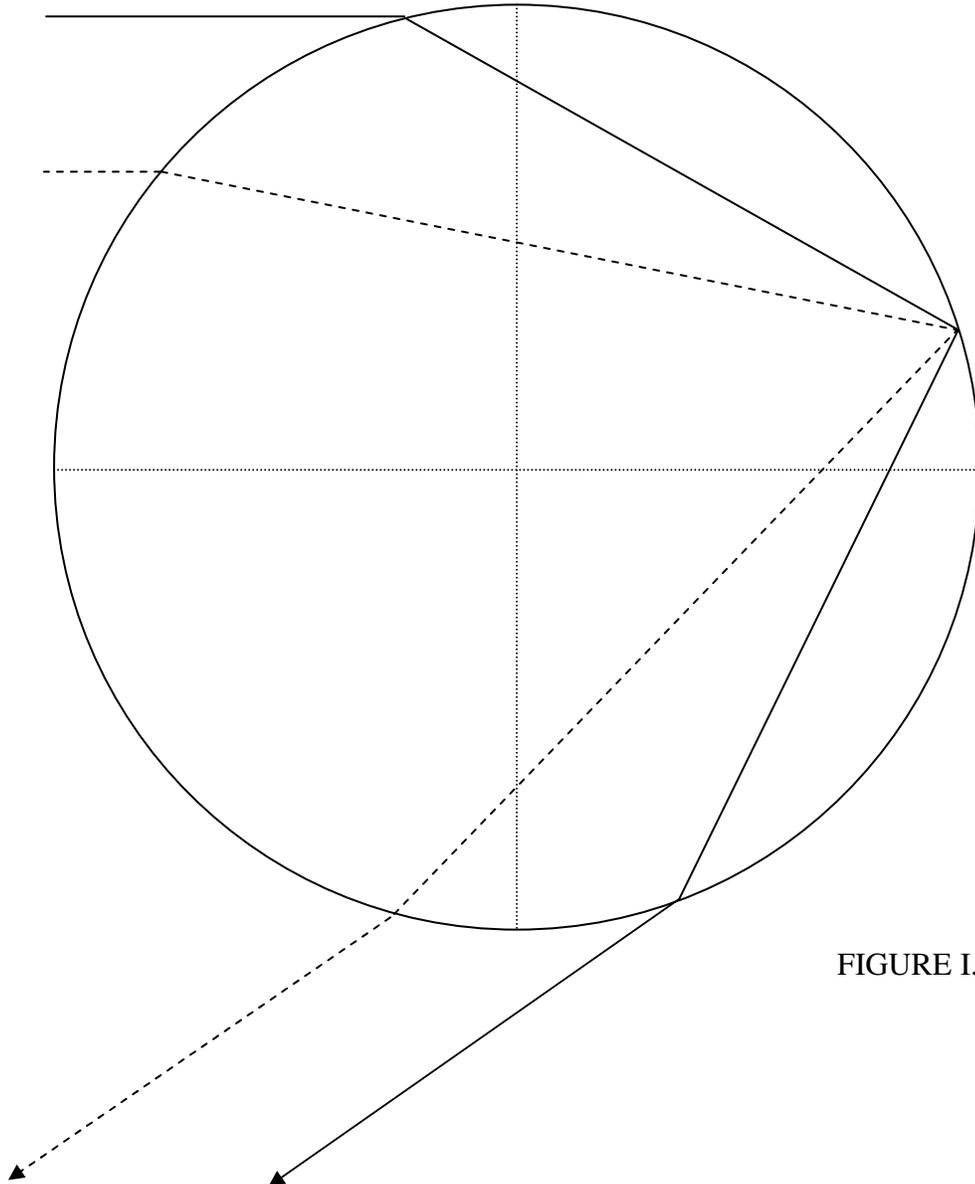


FIGURE I.18

If you look just inside the inner (blue) margin of the bow, you can often clearly see the interference fringes produced by two rays with the same deviation. I haven't tried, but if you were to look through a filter that transmits just one colour, these fringes (if they are bright enough to see) should be well defined. The optical path difference for a given deviation, or given r , depends on the radius of the drop (and on its refractive index). For a drop of radius a it is easy to see that the optical path difference is

$$2a(\cos \theta_2 - \cos \theta_1) - 4n(\cos \theta'_2 - \cos \theta'_1),$$

where θ_1 is the larger of the two angles of incidence. Presumably if you were to measure the fringe spacing, you could determine the size of the drops. Or, if you were to conduct a Fourier analysis of the visibility of the fringes, you could determine, at least in principle, the size distribution of the drops.

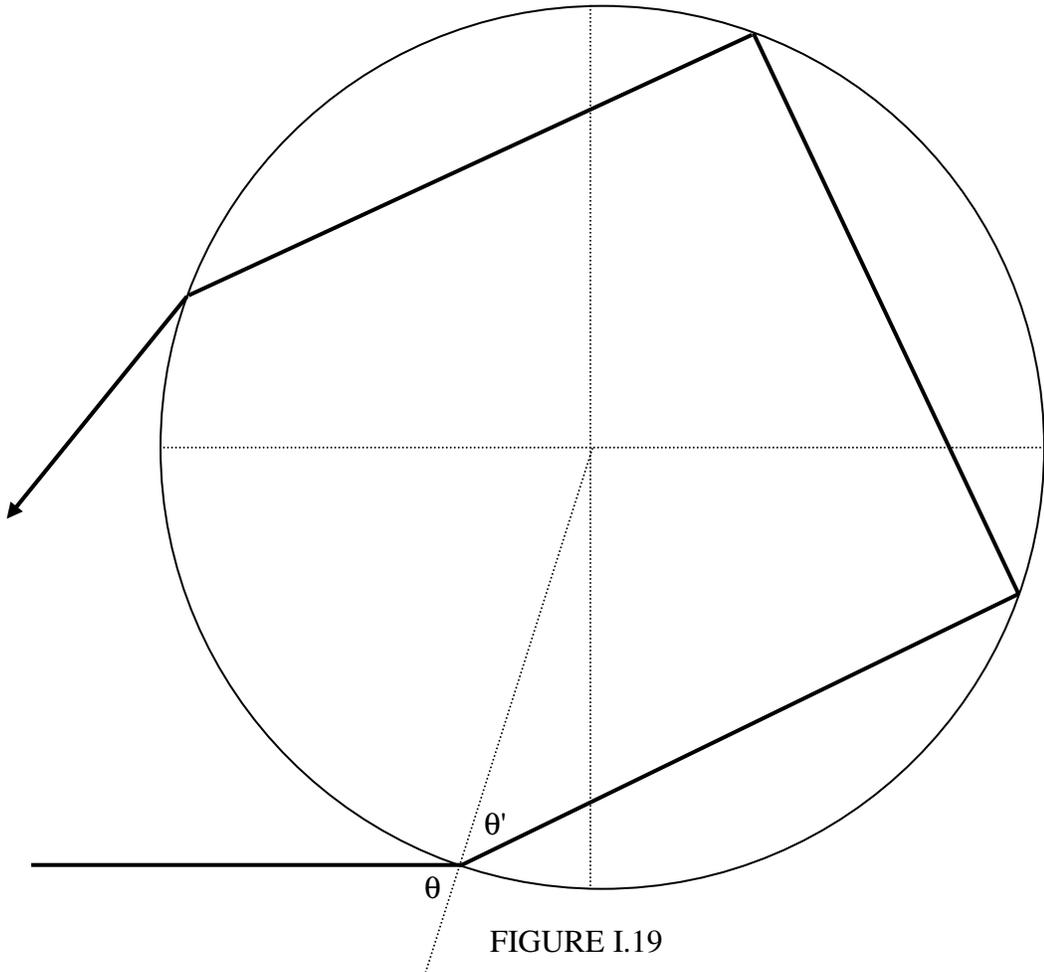
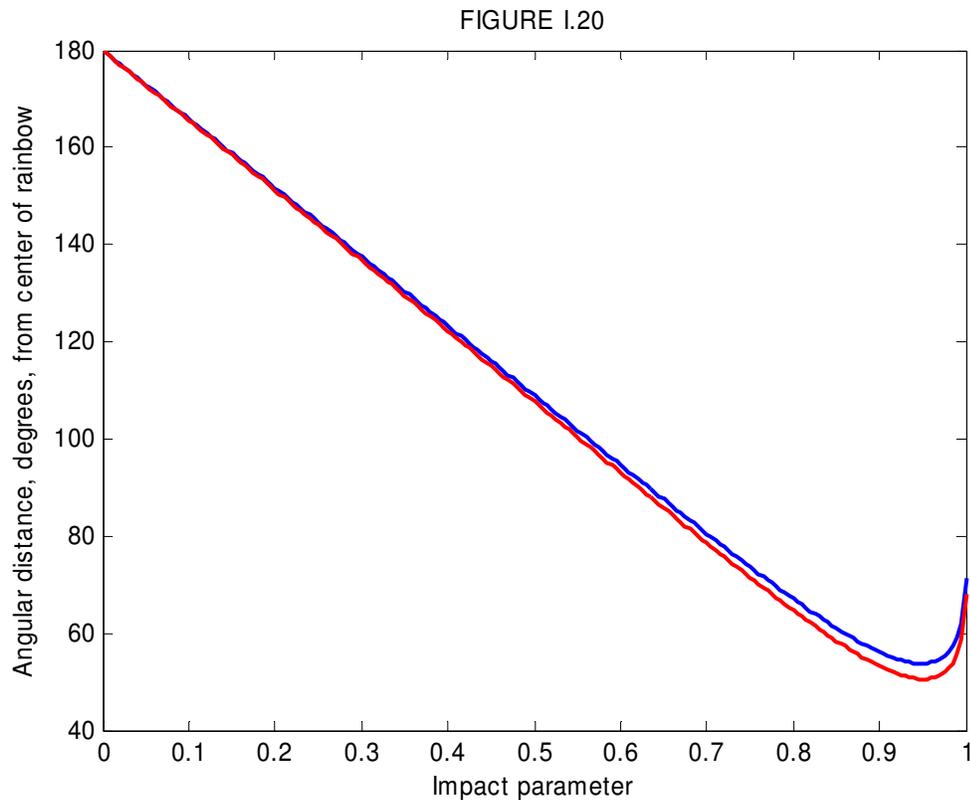


FIGURE I.19

Some distance outside the primary rainbow, there is a secondary rainbow, with colours reversed – i.e. red on the inside, blue on the outside. This is formed by two internal reflections inside the drop (figure I.19). The deviation of the final emergent ray from the direction of the incident ray is $(\theta - \theta') + (\pi - 2\theta') + (\pi - 2\theta') + (\theta - \theta')$, or $2\pi + 2\theta - 6\theta'$ counterclockwise, which amounts to $D = 6\theta' - 2\theta$ clockwise. That is,

$$D = 6\sin^{-1}(b/n) - 2\sin^{-1} b. \quad 1.7.9$$

clockwise, and, as before, this corresponds to an angular distance from the centre of the bow $r = \pi - D$. I show in figure I.20 the angular distance from the centre of the bow versus the impact parameter b . Notice that D goes through a *maximum* and hence r has a *minimum* value. There is no light scattered outside the primary bow, and no light scattered inside the secondary bow. When the full glory of a primary bow and a secondary bow is observed, it will be seen that the space between the two bows is relatively dark, whereas it is brighter inside the primary bow and outside the secondary bow.



Differentiation shows that the least value of r , (greatest deviation) corresponding to the radius of the secondary bow is

$$R = \pi - 6\sin^{-1}\sqrt{\frac{9-n^2}{8n^2}} - 2\sin^{-1}\sqrt{\frac{9-n^2}{8}}. \quad 1.7.10$$

For $n = 1.3439$ (blue) this is $53^\circ 42'$ and for $n = 1.3316$ (red) this is $50^\circ 31'$. Thus the red is on the inside of the bow, and blue on the outside.

Exercise. Trace a ray of light through a spherical drop with *three* internal reflections, leading to a *tertiary* rainbow. Determine an expression for the deviation as a function of impact parameter, and hence by calculus find the angle of minimum deviation. You should find that for “red” light ($n = 1.33$), the angle of minimum deviation is $42^\circ.85$ (impact parameter = 0.9740), and for “blue” light ($n = 1.34$), the angle of minimum deviation is $39^\circ.19$ (impact parameter = 0.9731). Since the angle of minimum deviation is less than 90° (i.e. *forward* scattering), this means that the tertiary rainbow will be centred on the Sun (not the anti-Sun), and the radius of the bow is equal to the angle of minimum deviation (not its supplement). Thus the tertiary bow has blue on the inside and red on the outside. The tertiary bow, and even a quaternary one, has been observed by Theusner (*Applied Optics* **50**, F129 (2011)). I am grateful to Grant Hutchison for drawing my attention to this paper.)

Let us return to the primary bow. The deviation is (equation 1.7.1) $D = \pi + 2\theta - 4\theta'$. Let's take $n = 4/3$, which it will be for somewhere in the middle of the spectrum. According to equation 1.7.8, the radius of the bow ($R = \pi - D_{\min}$) is then about 42° . That is, $2\theta' - \theta = 21^\circ$. If we combine this with Snell's law, $3\sin\theta = 4\sin\theta'$, we find that, at minimum deviation (i.e. where the primary bow is), $\theta = 60^\circ.5$ and $\theta' = 40^\circ.8$. Now, at the point of internal reflection, not all of the light is reflected (because θ' is less than the critical angle of $36^\circ.9$), and it will be seen that the angle between the reflected and refracted rays is $(180 - 60.6 - 40.8)$ degrees = $78^\circ.6$. Those readers who are familiar with Brewster's law will understand that when the reflected and transmitted rays are at right angles to each other, the reflected ray is completely plane polarized. The angle, as we have seen, is not 90° , but is $78^\circ.6$, but this is sufficiently close to the Brewster condition that the reflected light, while not completely plane polarized, is strongly polarized. Thus, as can be verified with a polarizing filter, the rainbow is strongly plane polarized.

I now want to address the question as to how the brightness of the bow varies from centre to circumference. It is brightest where the slope of the deviation versus impact parameter curve is least – i.e. at minimum deviation (for the primary bow) or maximum deviation (for the secondary bow). Indeed the radiance (surface brightness) at a given distance from the centre of the bow is (among other things) inversely proportional to the slope of that curve. The situation is complicated a little in that, for deviations between D_{\min} and $2\pi - 4\sin^{-1}(1/n)$, (this latter being the deviation for grazing incidence), there are two impact parameters giving rise to the same deviation, but for deviations greater than that

(i.e. closer to the centre of the bow) only one impact parameter corresponds to a given deviation.

Let us ask ourselves, for example, how bright is the bow at 35° from the centre (deviation 145°)? The deviation is related to impact parameter by equation 1.7.6. For $n = 4/3$, we find that the impact parameters for deviations of 144 , 145 and 146 degrees are as follows:

D°	b
144	0.6583 and 0.9623
145	0.6366 and 0.9693
146	0.6157 and 0.9736

Figure I.21 shows a raindrop seen from the direction of the approaching photons.

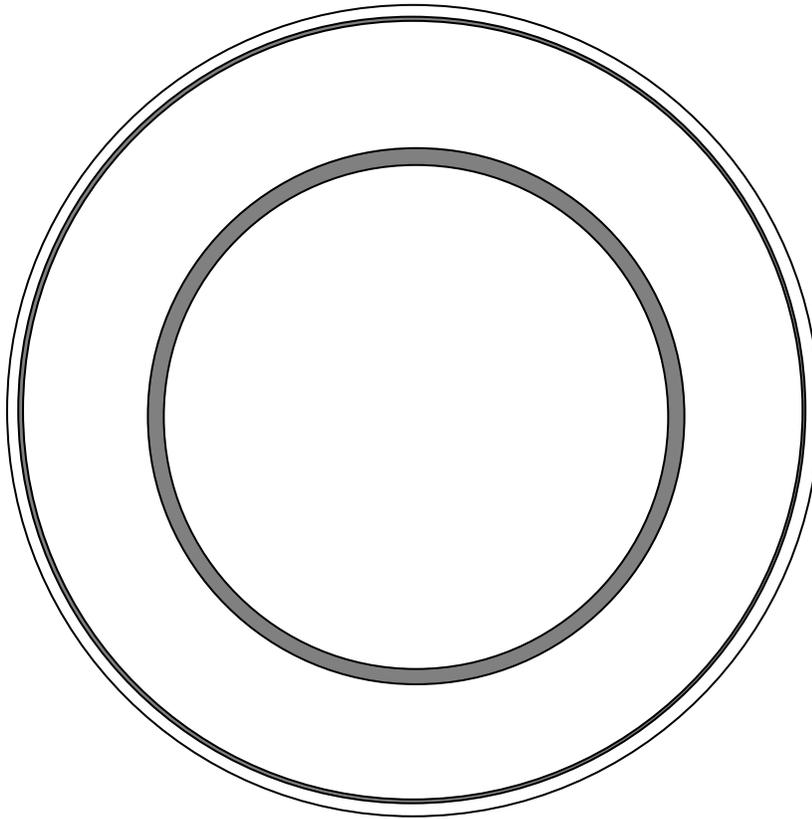
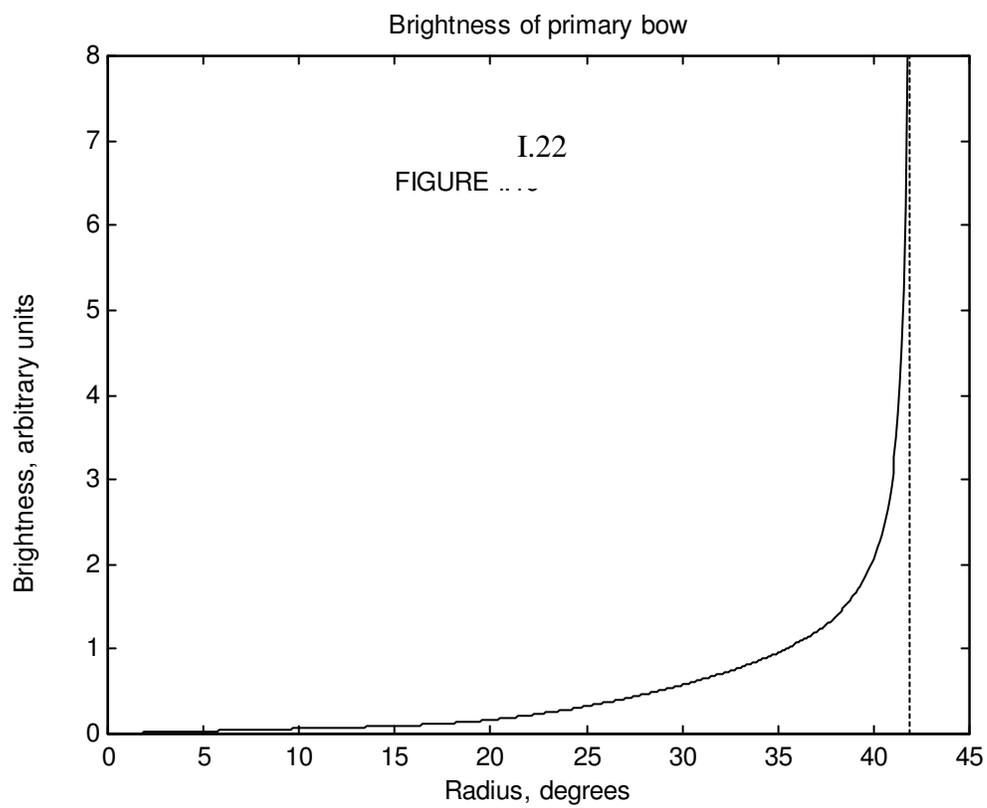


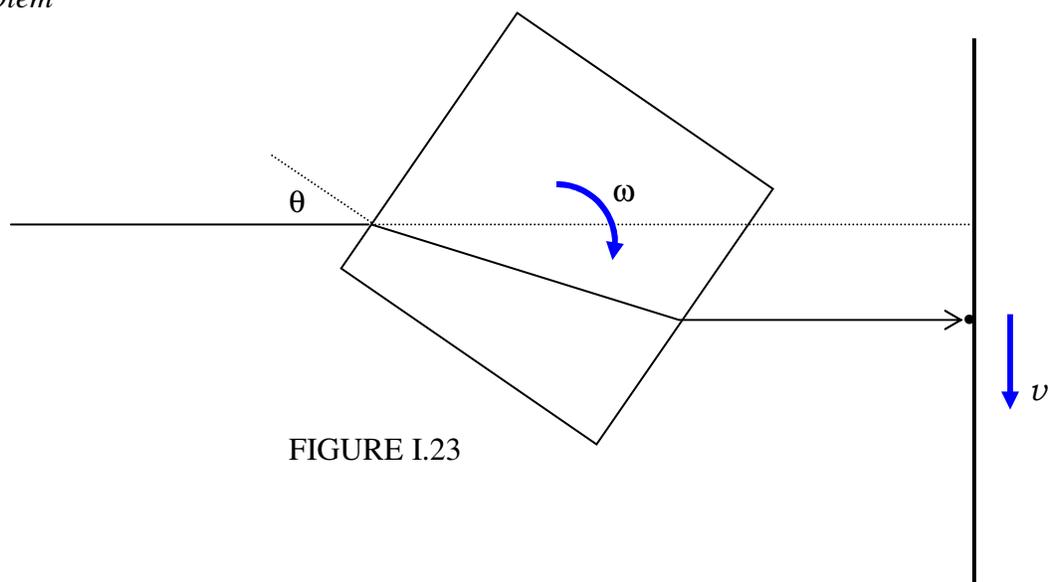
FIGURE I.21

Any photons with impact parameters within the two dark annuli will be deviated between 144° and 146° , and will ultimately approach the observer at angular distances between 36° and 34° from the centre. The radiance at a distance of 35° from the centre will be proportional, among other things, to the sum of the areas of these two annuli.

I have said “among other things”. Let us now think about other things. I have drawn figure I.15 as if all of the light is transmitted as it enters the drop, and then all of it is internally reflected within the drop, and finally all of it emerges when it leaves the drop. This is not so, of course. At entrance, at internal reflection and at emergence, some of the light is reflected and some is transmitted. The fractions that are reflected or transmitted depend on the angle of incidence, but, for minimum deviation, about 94% is transmitted on entry to and again at exit from the drop, but only about 6% is internally reflected. Also, after entry, internal reflection and exit, the percentage of polarization of the ray increases. The formulas for the reflection and transmission coefficients (Fresnel’s equations) are somewhat complicated (equations 1.5.1 and 1.5.2 are for unpolarized incident light), but I have followed them through as a function of impact parameter, and have also taken account of the sizes of the one or two annuli involved for each impact parameter, and I have consequently calculated the variation of surface brightness for one colour ($n = 4/3$) from the centre to the circumference of the bow. I omit the details of the calculations, since this chapter was originally planned as an elementary account of reflection and transmission, and we seem to have gone a little beyond that, but I show the results of the calculation in figure I.22. I have not, however, taken account of the interference phenomena, which can often be clearly seen just within the primary bow.



1.8 Problem



See figure I.23. A ray of light is directed at a glass cube of side a , refractive index n , eventually to form a spot on a screen beyond the cube. The cube is rotating at an angular speed ω . Show that, when the angle of incidence is θ , the speed of the spot on the screen is

$$v = a\omega \left(\cos \theta - \frac{n^2 \cos 2\theta + \sin^4 \theta}{(n^2 - \sin^2 \theta)^{3/2}} \right)$$

and that the greatest displacement of the spot on the screen from the undisplaced ray is

$$D = \frac{a}{\sqrt{2}} \left(1 - \frac{1}{\sqrt{n^2 - \sin^2 \theta}} \right).$$

I refrain from asking what is the maximum speed and for what value of θ does it occur. However, I ran the equation for the speed on the computer, with $n = 1.5$, and, if the formula is right, the speed is $\frac{1}{3}a\omega$ when $\theta = 0$, and it increases monotonically up to $\theta = 45^\circ$, which is as far as we can go for a cube. However, if we have a rectangular glass block, we can increase θ to 90° , at which time the speed is $0.8944 a\omega$. The speed goes through a maximum of about $0.9843a\omega$ when $\theta = 79^\circ.3$. I'd be interested if anyone can confirm this, and do it analytically.

1.9 Differential Form of Snell's Law

Snell's law in the form $n \sin \theta = \text{constant}$ is useful in calculating how a light ray is bent in travelling from one medium to another where there is a discrete change of refractive index. If there is a medium in which the refractive index is changing continuously, a differential form of Snell's law may be useful. This is obtained simply by differentiation of $n \sin \theta = \text{constant}$, to obtain the differential form of Snell's law:

$$\cot \theta d\theta = -\frac{dn}{n}. \quad 1.9.1$$

If we express this in terms of the complementary angle ψ (see figure I.24), this equation takes the form

$$\tan \psi d\psi = \frac{dn}{n}. \quad 1.9.2$$

Let us see how this might be used. Let us suppose, for example, that we have some medium in which the refractive index diminishes linearly with the y coordinate according to

$$n = n_0 - \frac{(n_0 - 1)y}{a}. \quad 1.9.3$$

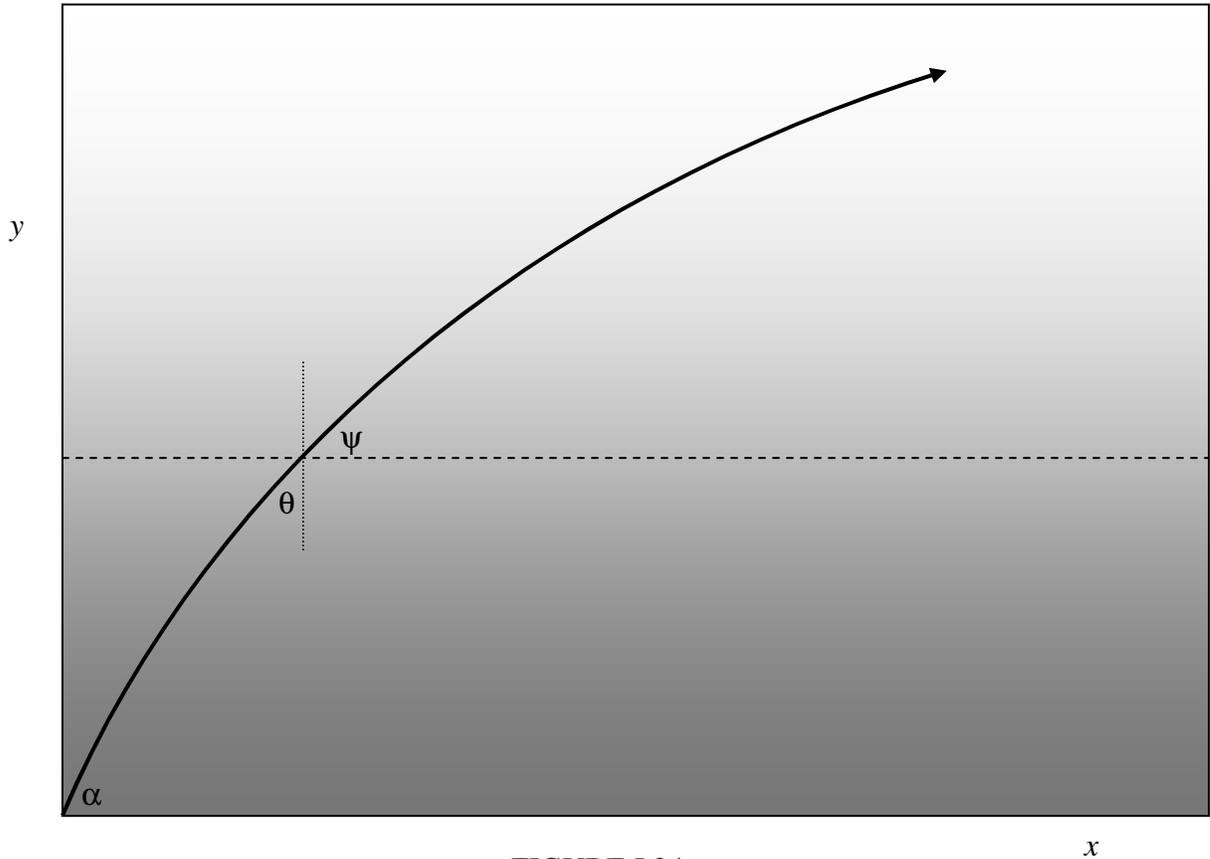


FIGURE I.24

This form has $n = n_0$ at $y = 0$, decreasing linearly to $n = 1$ at $y = a$. We suppose that $n = 1$ everywhere beyond $y = a$. The equation is simplified if we write μ for $n - 1$ and μ_0 for $n_0 - 1$, so that

$$\mu = \mu_0 \left(1 - \frac{y}{a} \right), \quad 1.9.4$$

which has μ decreasing linearly from μ_0 to zero.

Let us suppose that we direct a light ray upwards from the origin in a direction making an angle α with the horizontal, and we wish to trace the ray through the medium as the refractive index continuously changes. See figure I.24.

With the refractive index changing as in equation 1.9.4, Snell's law takes the form

$$\tan \psi \, d\psi = -\frac{dy}{k-y}, \quad \text{where } k = \frac{n_0 a}{\mu_0}. \quad 1.9.5$$

On integration, this becomes

$$y = k \left(1 - \frac{\cos \alpha}{\cos \psi} \right) \quad \text{or} \quad \cos \psi = \frac{k \cos \alpha}{k-y}. \quad 1.9.6$$

This is the (y, ψ) equation to the path taken by the light. One can see from this equation that the path becomes horizontal when $y = k(1 - \cos \alpha)$.

To find the (x, y) equation, we use the identities $\cos \psi = (1 + \tan^2 \psi)^{-1/2}$ and $\tan \psi = \frac{dy}{dx}$. Substitution of these into equation 1.9.6 gives the differential equation to the path in (x, y) coordinates:

$$\left(\frac{dy}{dx} \right)^2 = \frac{(k-y)^2 - l^2}{l^2}, \quad \text{where } l = k \cos \alpha. \quad 1.9.10$$

This can be readily integrated to give the (x, y) equation to the path:

$$x = l \ln \left[\frac{k(1 + \sin \alpha)}{k-y \pm \sqrt{(k-y)^2 - l^2}} \right]. \quad 1.9.11$$

In figure I.25 I have taken $a = 1$ (in other words, I am expressing all distances in units of a), and I am taking $n_0 = 1.5$ (hence $k = 3$), and I calculate the paths for $\alpha = 15^\circ, 30^\circ, 45^\circ, 60^\circ, 75^\circ$.

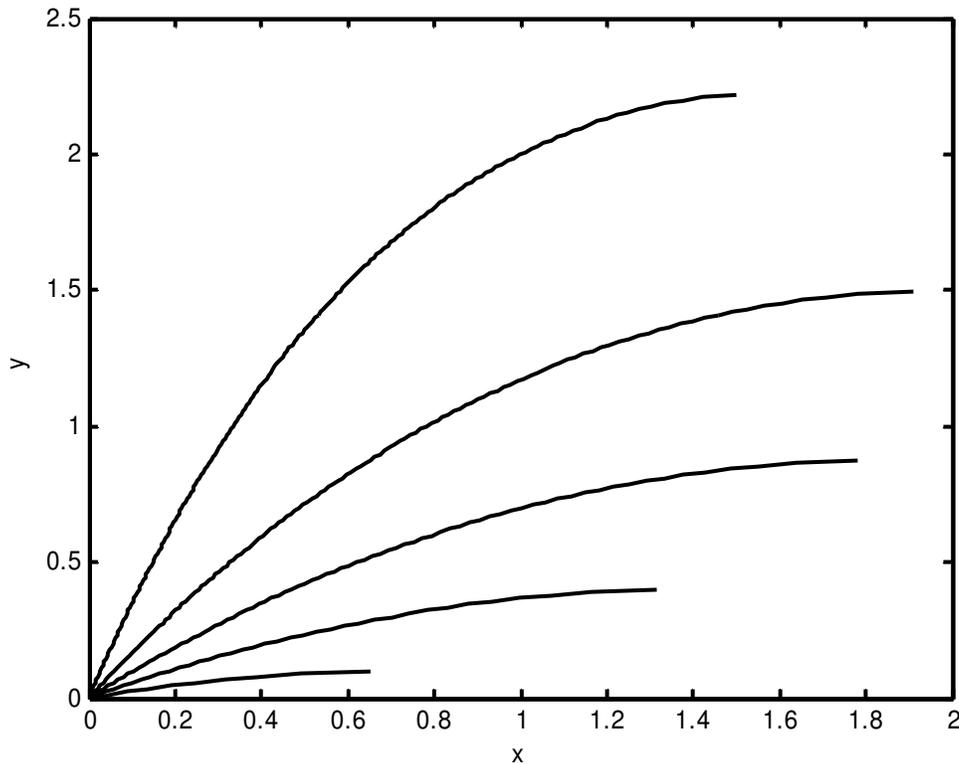


FIGURE I.25

Figure I.25 and equation 1.9.11 show the paths of the light if the refractive index is varying from n_0 at $y = 0$, to 1 at $y = a$ as described by equation 1.9.3. But now suppose that the refractive index is varying with height according to

$$n = \frac{n_0 a}{(n_0 - 1)y + a}. \quad 1.9.12$$

In this model, too, the refractive index goes from n_0 at $y = 0$, to 1 at $y = a$, but the variation is not linear. In fact you might like to convince yourself that it is the speed of light that is increasing linearly from $y = 0$ to $y = a$. See if you can trace the paths of the light rays in this situation. I think they are arcs of circles, and you might be able to calculate the radii and the coordinates of the centres of the circles.

Here's another one:

$$n = \sqrt{\frac{n_0^2 a}{(n_0^2 - 1)y + a}}. \quad 1.9.13$$

Here, too, the refractive index goes from n_0 at $y = 0$, to 1 at $y = a$. You might like to try tracing the rays in this model. I think they may be arcs of cycloids.

Of course these examples may seem to be very unlikely. Can you imagine a glass block of width a , made of glass whose refractive index varies continuously from 1.5 at one edge to zero at the other? Not very likely, yet there is a situation that comes to mind in which there is a continuous variation of refractive index from some basal value n_0 to zero. I am thinking of Earth's atmosphere (or indeed the atmosphere of any planet). As light from a star travels through Earth's atmosphere, it moves not in a straight line, but in a slight curve, so that it is deviated through a few arc minutes before it reaches the astronomer's telescope. For a star low down near the horizon, the refraction amounts to almost half a degree. This has to be taken into account when astronomers are making precise positional measurements. And sound waves, passing through the atmosphere, are also subject to refraction via the differential form of Snell's law. The speed of sound (and hence the refractive index) varies with the temperature of the atmosphere, and hence with height in the atmosphere. Among the many applications of this sort of theory is the path of sound waves from meteorites hurtling through the atmosphere. This is discussed in a paper published in *Meteoritics and Planetary Science* **34**, 572-585 (1999).