## Module P6.2 Rays and geometrical optics

1 Opening items
1.1 Module introduction
1.2 Fast track questions
1.3 Ready to study?

2 Waves, rays and geometrical optics
2.1 Rays and rectilinear propagation
2.2 Waves and the ray approximation

3 Reflection
3.1 The law of reflection
3.2 Image formation in a plane mirror

4 Refraction
4.1 The law of refraction
4.2 Image formation by a plane interface
4.3 Total internal reflection and optical fibres
4.4 Continuous refraction and atmospheric optics

5 Fermat's principle
5.1 Reflection at a plane mirror
5.2 Refraction at a plane interface
5.3 The principle of reversibility
5.4 The physical basis of Fermat's principle

6 Closing items
6.1 Module summary
6.2 Achievements
6.3 Exit test

## 1 Opening items

### 1.1 Module introduction

From the earliest period of civilization people have pondered over their reflections in pools of water and contemplated the distorted images of objects under the surface of those pools. Such everyday phenomena have intrigued and challenged natural philosophers for millennia, yet it was not until the end of the 16th century that the simple rules that govern them were presented in something akin to their modern mathematical form. These rules, all of which assume that light in a uniform medium travels along linear paths called rays, now provide the basis of geometrical optics, a field of study in which many optical conundrums are reduced to problems in geometry and trigonometry.
This module provides an introduction to geometrical optics. It describes the conditions under which the behaviour of light may be adequately described in terms of rays and uses the ray concept to describe reflection at a plane mirror (Section 3), refraction at a plane interface (Subsection 4.2), total internal reflection in a medium of relatively high refractive index and continuous refraction in a medium where the refractive index changes from place to place (Subsections 4.3 and 4.4). The phenomenon of optical mirages will be described as will the passage of light through an optical fibre (Subsection 4.4).
The module closes with a discussion (in Section 5) of Fermat's principle which provides a unified basis for geometrical optics from which all of the subject's basic laws and principles can be derived.
FLAP P6.2 Rays and geometrical optics

COPYRIGHT © 1998 THE OPEN UNIVERSITY
S570 V1.1
$\Theta($

Most of the basic ideas in geometrical optics have been known for almost four hundred years, but do not fall into the trap of thinking that geometrical optics is a dead subject, of little relevance to the modern world - far from it. The very latest designs for camera lenses or telescope mirrors depend on intelligent application of the laws of optics, now augmented by computer power. Similarly, modern communication networks use the latest developments in optical fibre technology and rely on these same ideas.

Study comment Having read the introduction you may feel that you are already familiar with the material covered by this module and that you do not need to study it. If so, try the Fast track questions given in Subsection 1.2. If not, proceed directly to Ready to study? in Subsection 1.3.
$\longrightarrow \leq$

### 1.2 Fast track questions

Study comment Can you answer the following Fast track questions?. If you answer the questions successfully you need only glance through the module before looking at the Module summary (Subsection 6.1) and the Achievements listed in Subsection 6.2. If you are sure that you can meet each of these achievements, try the Exit test in Subsection 6.3. If you have difficulty with only one or two of the questions you should follow the guidance given in the answers and read the relevant parts of the module. However, if you have difficulty with more than two of the Exit questions you are strongly advised to study the whole module.

## Question F1

What is meant by the ray approximation in optics? Under what conditions can one use the ray approximation to describe the transmission of light?

## Question F2

Using the laws of reflection of light, draw a diagram to show how your eye can see the virtual image of a point object in a plane mirror. If the object is 20 cm from the mirror, where is the image?

## Question F3

(a) A light ray inside a glass block (refractive index 1.50) strikes a glass-air interface at an angle $\theta_{1}$ to the normal. Find the critical angle $\theta_{1 \mathrm{c}}$ beyond which the ray is totally internally reflected.
(Take the refractive index of air to be 1.00.)
(b) Explain how total internal reflection is used in optical fibres.

Study comment Having seen the Fast track questions you may feel that it would be wiser to follow the normal route through the module and to proceed directly to Ready to study? in Subsection 1.3.

Alternatively, you may still be sufficiently comfortable with the material covered by the module to proceed directly to the Closing items.

### 1.3 Ready to study?

Study comment In order to study this module you will need to be familiar with certain ideas concerning the wave model of light. In particular you should know that many of the properties of light can be accounted for by assuming that light travels as waves characterized by a wavelength $\lambda$, a frequency $f$ and a speed $v$ that are related by $v=f \lambda$; that different colours of light correspond to different frequencies or wavelengths, and that the phenomena of interference and diffraction provide evidence in support of the wave model. It is not assumed that you have a detailed knowledge of these topics, indeed most of them are given a brief introduction when they first arise. However, if you are completely unfamiliar with these ideas you may find the introductions too brief, in which case you should refer to the Glossary which will indicate where they are more fully developed. The mathematical requirements of this module include trigonometry, and the geometry of triangles. In Section 5 some use will be made of differentiation to determine the stationary points (in this case local minima) of a function, but if this technique is unfamiliar to you then you should nevertheless still be able to meet the achievements of the module. If you are unsure about any of these items you can review them now by reference to the Glossary, which will also indicate where in FLAP they are developed. The following Ready to study questions will allow you to establish whether you need to review some of the topics before embarking on this module.

2

## Question R1

$\theta$ is an angle between $0^{\circ}$ and $90^{\circ}$ satisfying the equation $\sin \theta=0.45$. Find $\theta$.

## Question R2

Given the approximate values for the wavelengths in a vacuum of blue and red light as 450 nm and 650 nm , respectively, calculate the frequency of the light in each case given that the speed of light in a vacuum is $3.00 \times 10^{8} \mathrm{~m} \mathrm{~s}^{-1}$. (Note that $1 \mathrm{~nm}=10^{-9} \mathrm{~m}$.)


## Question R3

What is meant by saying two waves are in phase at a certain point? What is meant by saying two waves arriving at a common point by different routes interfere when they meet?

## 2 Waves, rays and geometrical optics

### 2.1 Rays and rectilinear propagation

'Light,' it is often said, 'travels in straight lines.' Of course, this is an oversimplification; lenses and mirrors can alter the course of a light beam, as can changes or irregularities in the medium (i.e. material) through which the light is moving. Nonetheless, the idea that light travels in straight lines in a uniform medium, sometimes referred to as the principle of rectilinear propagation, does provide a simple characterization of the behaviour of light under a wide range of circumstances. $\frac{1989}{}$ Evidence in support of the principle is provided by the 'shafts' of sunlight that can sometimes be seen streaming through open windows or through gaps in clouds and by countless other natural phenomena. The effect can easily be demonstrated in the laboratory by letting light from a very small source pass through an aperture (cut in an otherwise opaque piece of card) and fall on a screen. This is shown in Figure 1, where the aperture has a rectangular shape. The precise


Figure 1 Light from a small source passes through a rectangular aperture to form a well defined bright patch on a screen. shape of the bright patch on the screen can be predicted by drawing straight lines from the source to the screen. Any line hitting the opaque part of the card is stopped, otherwise it passes through to the screen.
FLAP P6.2 Rays and geometrical optics
$\begin{array}{llll}\text { COPYRIGHT © } 1998 & \text { THE OPEN UNIVERSITY } & \text { S570 V1.1 }\end{array}$

The idea that light travels in straight lines leads naturally to the concept of a ray of light.
A ray is a directed line (i.e. a line with an arrow on it) showing a possible path along which light can travel and indicating the direction in which such light would transport energy.

The principle of rectilinear propagation leads us to expect that light rays in a uniform medium will be straight lines (this is sometimes referred to as a pencil of rays), like those shown in Figure 1, though we know that rays may be kinked or reversed at mirrors or at the boundaries between different transparent media, and we should also expect rays to be bent or curved when the optical properties of a medium change from place to place.

### 2.2 Waves and the ray approximation

Although the idea of rectilinear propagation is intuitively appealing and accords with much of our everyday experience, it is still not fully consistent with all the observed properties of light. It has been known at least from the time of Isaac Newton (1642-1727) that when light travels through a very small aperture or passes close to a very small obstacle it spreads out in a way that cannot be simply explained in terms of motion in straight lines. For example, if the rectangular aperture in Figure 1 is less than about 0.1 mm wide then the bright patch on a screen placed 1 m away will not be perfectly sharp, even for the smallest possible source. Moreover, the blurring of the bright patch will become progressively greater if the size of the aperture is made even smaller. The implication of this is that, in this situation, light does not travel along straight lines and the concept of a light ray is not useful.


Figure 1 Light from a small source passes through a rectangular aperture to form a well defined bright patch on a screen.

The tendency of light to spread out when passing through narrow apertures is one manifestation of a widespread phenomenon known as diffraction. In general diffraction is characteristic of wave phenomena, such as water waves on a pond or sound waves in air. It arises because of the natural tendency of waves to spread out when they encounter obstacles or pass through apertures comparable in size to the wavelength of the waves. The diffraction of light by narrow slits is evidence that light too is a wave phenomenon and that the relevant waves have a very small wavelength.

The wave model of light, according to which light travels in the form of waves, with different wavelengths corresponding to different colours, is discussed more fully elsewhere in FLAP. Although we will mention it from time to time, we will not pursue it in any detail in this module. However, it should be noted that the use of rays to indicate possible pathways for light is entirely consistent with the wave model. In the context of the wave model, light of a given colour spreading out from a point-like source can be viewed as roughly analogous to the concentric ripples that spread out from a point of disturbance on the surface of a pond. Each expanding circular ripple can be thought of as representing an expanding spherical wavefront $\frac{1 \text { 国 }}{\text { s }}$ of the light, and the unimpeded movement of such wavefronts through a uniform medium can be indicated at any point by drawing a suitably directed ray perpendicular to the wavefront at that point.

The same principle, that rays are perpendicular to wavefronts in a uniform medium (see Figure 2), can be applied in other situations, provided diffraction effects are negligible. Thus, rays provide a useful way of describing the propagation of waves. Of course, the applicability of the ray description is limited by diffraction, so it is customary to speak of the ray approximation - an approximation that should be expected to break down when light passes through a narrow aperture or in any other situation where diffraction cannot be neglected.

Despite this limitation, there are many problems involving light that can be adequately treated by considering rays. These problems fall within the domain of geometrical optics, the field of study in which the ray approximation is valid. This module provides an introduction to some of the principles and applications of geometrical optics.


Figure 2 In terms of the wave model of light, rays are drawn perpendicular to wavefronts at any point and indicate the direction in which those wavefronts are moving.

## 3 Reflection

An important topic in geometrical optics is the reflection of light. In reflection, light strikes a surface and is returned to the medium from which it came. A surface which reflects strongly is said to have a high reflectivity and to be a good reflector. Such a reflector is commonly called a mirror. The simplest type of mirror, and the only type we will consider in this module, is a flat or plane mirror. A plane mirror can be made by coating a smooth flat surface (e.g. glass) with a metal film (e.g. silver) of thickness greater than a few wavelengths of light. Light which strikes the mirror cannot penetrate far into the metal and is almost entirely reflected. Some light however will be absorbed by the metal film and the extent of this absorption will generally depend on the wavelength of the incident light. For example, silver and aluminium are particularly good for mirrors because their absorption is almost independent of wavelength; in contrast, gold selectively absorbs the shorter visible wavelengths which causes a gold mirror to have its characteristic colour when viewed in ordinary daylight.

### 3.1 The law of reflection

Figure 3 shows a light ray from any point $P$ striking a mirror at point $A$, which is called the point of incidence, and, after reflection, passing through a point $\mathrm{P}^{\prime}$. The line drawn perpendicular to the mirror at A is called the normal and the angles made by the incoming incident ray and the outgoing reflected ray to the normal at the point of incidence are known as the angle of incidence $\theta_{\mathrm{i}}$ and the angle of reflection $\theta_{\mathrm{R}}$, respectively. Experiments clearly demonstrate the following law of reflection, which has been known since ancient times.

## Law of reflection

- The reflected ray, the incident ray and the normal all lie in the same plane.
- The angle of reflection is equal to the angle of incidence: $\theta_{\mathrm{R}}=\theta_{\mathrm{i}}$


Figure 3 A ray of light reflected from a plane mirror.

The two parts of this law can be derived from the wave model of light but here we take them as being justified by experimental observations. Notice that in Figure 3 the direction of the rays is reversible, that is, a ray from $\mathrm{P}^{\prime}$ to A will be reflected in such a way that it passes through P. This is a particular example of another general principle of geometrical optics to which we will return later.

## Question T1

Two plane mirrors are fixed together at right angles in the form of a letter 'L' with the reflecting surfaces on the inside. Show that any ray of light striking both mirrors will be reflected back in a direction parallel to its original direction.


Figure 3 A ray of light reflected from a plane mirror.

### 3.2 Image formation in a plane mirror

Mirrors and lenses are often used to make a bundle of light rays that diverge (i.e. spread out) from one point come together (i.e. converge) at some other point. This process is illustrated in Figure 4 where some of the rays diverging from a point object at O are reflected from a plane mirror and then brought together at a point I by a human eye. Under such circumstances we say that there is a real image of the object at the point I.

Figure 4 A bundle of rays diverging from O and reflected by a plane mirror, forms a virtual image at $I^{\prime}$ behind the mirror, while the human eye forms a real image at I.


If the eye shown in Figure 4 was your own, and you were asked where the object O appeared to be you would probably say that you could see it 'in' the mirror. The reason for this is fairly clear: the reflected rays travel in such a way that if it is assumed they always travel in straight lines then they must have come from the point $\mathrm{I}^{\prime}$ behind the mirror. Because the rays do not really diverge from this point but only appear to do so, we say that there is a virtual image of $O$ at the point $I^{\prime}$.

Note that this virtual image at $\mathrm{I}^{\prime}$ exists even when there is no eye to form the real image at I : it is sufficient that the reflected rays can be continued back to the point $\mathrm{I}^{\prime}$ 。 $\underline{\underline{2} 88}$

Bearing in mind the law of reflection, which tells us that the angles of incidence and reflection must be equal for each of the rays in Figure 4, it is fairly easy to demonstrate that the virtual image at $I^{\prime}$ is as far behind the mirror surface as the object at O is in front of it.

Figure 4 A bundle of rays diverging from O and reflected by a plane mirror, forms a virtual image at $I^{\prime}$ behind the mirror, while the human eye forms a real image at I.


## Question T2

Show that all rays coming from $O$ that are reflected by the mirror in Figure 4 will, after reflection, behave as though they came from a single fixed point $\mathrm{I}^{\prime}$.
(Hint: Consider one ray along the normal from O to the mirror and a general ray at an angle $\theta_{\mathrm{i}}$ to the normal.) Prove that the virtual image at $\mathrm{I}^{\prime}$ is at the same distance from the mirror as the object at O .

Figure 4 A bundle of rays diverging from O and reflected by a plane mirror, forms a virtual image at I' behind the mirror, while the human eye forms a real image at I.


We can formalize the result of Question T 2 by associating a Cartesian coordinate system with the mirror, as shown in Figure 5, with the $x$ - and $y$-axes in the plane of the mirror and the $z$-axis normal to the mirror. If the coordinates of the object at point O are $(X, Y, Z)$ then the coordinates of the point N (where the normal from O meets the mirror) will be ( $X, Y, 0$ ), and since the length $\mathrm{NI}^{\prime}$ is $Z$, the coordinates of the virtual image point $I^{\prime}$ are $(X, Y,-Z)$. Thus, when a plane mirror is used to form a virtual image of a point object, the coordinates of the image may be found from the coordinates of the object by changing the sign of the coordinate measured normal to the mirror while leaving the coordinates measured parallel to the mirror unchanged.

Now imagine looking at yourself in a mirror. Better still, try it! Your body is an example of an extended object rather than a point object, but the principle outlined above still applies. Each point on your body will correspond to a unique virtual image point behind the mirror, thus creating a virtual image of yourself which will be the same size as


Figure 5 A coordinate system fixed in a mirror. (The mirror is in the $(x, y)$ plane.) the original but which will appear to differ in one important respect. If you wave your right hand your mirror image will wave its left hand, and if you shut your left eye your mirror image will shut its right eye. We describe this by saying that the image is reversed, i.e. left and right appear to have been interchanged.

At the same time, the image is said to be erect, i.e. the right way up - the mirror does not appear to interchange your head and your feet. The ability of a mirror to produce an image that is simultaneously reversed and yet erect has been used to perplex students for generations and you are not going to be spared that perplexity!
In the Exit test at the end of this module you will be asked to explain how a plane mirror can apparently interchange left and right without interchanging top and bottom. You might like to keep that in mind as you read the next paragraph; it won't tell you the answer, but it may help you to formulate one.


Figure 5 A coordinate system fixed in a mirror. (The mirror is in the ( $x, y$ ) plane.)

We can gain more insight into the nature of the reflected image by supposing that the origin of a coordinate system is located at some point in front of a plane mirror, as shown in Figure 6, so that the system itself can be seen reflected in the mirror. How does the virtual image formed by the mirror differ from the real coordinate system? As indicated in Figure 6 , the $x$ - and $y$-axes of the image are parallel to those of the real system, but the $z$-axis is reversed. This might seem a rather insignificant change - you might think that by rotating the real system you could reorientate it so that all of its axes are parallel to the corresponding axes of the image, but if so, you would be wrong. The real system and its mirror image are fundamentally different in the same sense that your right hand is fundamentally different from your left hand; no matter how you rotate your right hand it never becomes a left hand, and no matter how you rotate the real coordinate system in Figure 6 you will never completely align it with its mirror image.


Figure 6 A right-handed coordinate system, reflected in a plane mirror, forms an image which is a left-handed system. (The mirror is in the $(x, y)$ plane.)

In fact the real system shown in Figure 6 is called a right-handed coordinate system because if you imagine yourself at the origin of the coordinate system, looking in the direction of increasing $z$, it requires a right-handed (clockwise) screw rotation to go from the $x$-axis to the $y$ axis, whereas the image system is called a left-handed coordinate system because it requires a left-handed (anticlockwise) screw rotation to go from the $x^{\prime}$-axis to the $y^{\prime}$-axis if you are looking in the direction of increasing $z^{\prime}$.

No rotation will ever succeed in aligning a right-handed coordinate system with a left-handed system nor a left-handed system with a righthanded system. However, the mirror image of a right-handed system is a left-handed system which may be rotated into alignment with any other left-handed system; and the mirror image of a left-handed system is a right-handed system which may be rotated into alignment with any other right-handed system. This interchange of right-handed and left-handed coordinate systems is the essence of image reversal.


Figure 6 A right-handed coordinate system, reflected in a plane mirror, forms an image which is a left-handed system. (The mirror is in the $(x, y)$ plane.)

To summarize:

When light from an object is reflected in a plane mirror a reversed virtual image is formed. This image is erect, the same size as the original object and located as far behind the mirror as the object is in front.

Although we have described the formation of images in plane mirrors produced by coating a glass surface with a metallic film, you will undoubtedly have seen reflections from other surfaces such as polished bench tops or unsilvered glass. Any surface acts to some extent as a reflector, though its reflectivity generally depends on the angle of incidence and increases as that angle approaches $90^{\circ}$. This phenomenon can be explained by the wave model of light but not in terms of rays and geometrical optics. Nonetheless, it is interesting to note that mirrors can be made from some rather unlikely materials, provided we restrict the angle of incidence to be nearly $90^{\circ}$ - a condition usually referred to as grazing incidence.

For example, a piece of paper is a very poor reflector of light at or near normal incidence, but at grazing incidence it can actually be quite good. Try this for yourself - hold a piece of paper up so that you can look along its surface at a light bulb and you should see a surprisingly bright reflection of the bulb. The quality of this image will not be very good because the surface of the paper will not be very flat when measured on the scale of the wavelength of light, still, it is quite surprising that paper can perform this function at all.
FLAP P6.2 Rays and geometrical optics
$\begin{array}{llll}\text { COPYRIGHT © } 1998 & \text { THE OPEN UNIVERSITY } & \text { S570 V1.1 }\end{array}$

## 4 Refraction

If you stand by a window in a brightly lit room at night, with the curtains open, you can usually see your own reflection in the window while someone outside can see you through the window. Your ability to see yourself in a 'transparent' window demonstrates the phenomenon of partial reflection. When light strikes the interface (i.e. boundary) between two different materials some light is usually reflected back into the first material while the rest is transmitted or absorbed by the second material. If both the materials involved are transparent then relatively little energy will be absorbed and any incident ray will split into two at the interface, one part being reflected while the other is transmitted. The light entering the second material is said to undergo refraction at the interface, which may cause it to travel in a different direction from the incident light. The proportions reflected and refracted at the interface will depend on the nature of the two materials and the angle of incidence, and will generally vary with wavelength.

### 4.1 The law of refraction

Figure 7 shows a single ray of light travelling through a transparent medium (material 1) and striking the interface with a different transparent medium (material 2). The ray is partly reflected and the reflected ray obeys the usual law of reflection. The remaining part of the ray is refracted at the interface and then travels through the second material; this is called the refracted ray. The direction of the refracted ray is described by the angle of refraction $\theta_{\mathrm{r}}$, which is the angle between the refracted ray and the normal (in material 2) to the interface at the point of incidence.


Figure 7 Reflection and refraction of a ray of light at the interface between two different transparent materials, where $\mu_{2}>\mu_{1}$.

Extensive experimentation with many kinds of transparent materials supports the following law of refraction.

## Law of refraction

- The incident ray, the refracted ray and the normal all lie in the same plane.
- The angle of incidence and the angle of refraction are related by

$$
\begin{equation*}
\frac{\sin \theta_{\mathrm{i}}}{\sin \theta_{\mathrm{r}}}=\text { constant } \tag{1a}
\end{equation*}
$$

where the constant depends on the two materials involved and on the wavelength of the light. $\underline{4}$


Figure 7 Reflection and refraction of a ray of light at the interface between two different transparent materials, where $\mu_{2}>\mu_{1}$.

This law was first propounded in 1621 by the Dutch mathematician and astronomer Willebrord Snell (1591-1626) and was published after his death by the French philosopher René Descartes (1596-1650). Equation 1a

$$
\frac{\sin \theta_{\mathrm{i}}}{\sin \theta_{\mathrm{r}}}=\text { constant }
$$

## (Eqn 1a)

is usually referred to as Snell's law, though that term is also applied to a more detailed form of the law that we will now introduce.

As in the case of reflection, the directions of the incident and refracted rays in Figure 7 are reversible, that is, a ray incident at an angle $\theta_{\mathrm{r}}$ in material 2 will be refracted at an angle $\theta_{\mathrm{i}}$ in material 1. Because of this it makes good sense to replace $\theta_{\mathrm{i}}$ and $\theta_{\mathrm{r}}$ in Equation 1a by $\theta_{1}$ and $\theta_{2}$ since the numerical subscripts show the material in which each angle is to be measured but don't imply that either is necessarily an angle of incidence or refraction. Moreover, experience shows that the constant


Figure 7 Reflection and refraction of a ray of light at the interface between two different transparent materials, where $\mu_{2}>\mu_{1}$. in Equation 1a can be written as the quotient $\mu_{2} / \mu_{1}$, where $\mu_{1}$ is a property of material 1 , called the refractive index of material 1 , and $\mu_{2}$ the refractive index of material 2.

Taking these points into account we obtain a more informative version of Equation 1a:

$$
\frac{\sin \theta_{\mathrm{i}}}{\sin \theta_{\mathrm{r}}}=\text { constant }
$$

(Eqn 1a)

Snell's law $\frac{\sin \theta_{1}}{\sin \theta_{2}}=\frac{\mu_{2}}{\mu_{1}}$

## Question T3

A ray of light passes through a parallel-sided block of glass surrounded by air. Show that the ray leaving the block is parallel to the ray entering it.

For light of a given wavelength, the refractive index of a medium is a dimensionless quantity that depends on the physical properties of that medium. Table 1 gives the refractive indices of a variety of transparent materials for light of one particular wavelength. Media of high (or low) refractive index are said to have a high (or low) optical density. It is often useful to identify which of the two media at an interface is the more optically dense since this will determine whether or not rays entering that material will be refracted towards the normal or away from it.

If light passes from material A to material $B$, where material $B$ is more optically dense than material $A$, in which direction will the light be deflected? (Assume $\theta_{\mathrm{i}} \neq 0$.)

Table 1 The refractive indices of some transparent materials.

| Material | Refractive index |
| :---: | :---: |
| vacuum | 1.00 (exactly) |
| air* | 1.00029 |
| water | 1.33 |
| crown glass | 1.52 |
| fused quartz | 1.46 |
| diamond | 2.42 |

The refractive index of a medium also determines the speed of light in that medium. If we let $c$ represent the speed of light in a vacuum ( $c=3.00 \times 10^{8} \mathrm{~m} \mathrm{~s}^{-1}$, to three significant figures) and $v$ the speed of light in a medium, then the refractive index $\mu$ of that medium is given by:

$$
\begin{equation*}
\mu=\frac{c}{v} \tag{2}
\end{equation*}
$$

Applying this relationship to both of the media at an interface allows us to rewrite Snell's law in terms of $v_{1}$ and $v_{2}$, the respective speeds of light in materials 1 and 2 :

$$
\begin{equation*}
\frac{\sin \theta_{1}}{\sin \theta_{2}}=\frac{v_{1}}{v_{2}} \tag{1c}
\end{equation*}
$$

This equation not only shows that a light ray is generally refracted towards the normal when it enters an optically denser medium (if $v_{2}$ is less than $v_{1}$, then $\theta_{2}$ must be less than $\theta_{1}$ ), it also gives some physical insight into why this should be so.

The situation is analogous to that of a car driving across a field that suddenly encounters a patch of mud. Whichever wheel enters the mud first will be slowed down so that the car as a whole tends to turn around that wheel until both wheels are in the mud and both wheels are moving forward at the same speed to each other. Once the speed of the wheels has been matched the car will proceed in a straight line, albeit at a somewhat reduced rate. A similar argument can be applied in reverse to explain why light entering a less optically dense medium tends to be refracted away from the normal as its speed increases.

It has already been mentioned that the refractive index of a material depends on the wavelength of the light involved. The nature of this dependence for one form of glass, fused quartz, is shown in Figure 8. The variation in $\mu$ is not very great, just $1 \%$ or so over the full range of visible wavelengths, but it is sufficient to have easily observable consequences. Because $\mu$ varies with wavelength, the extent to which a ray of light is refracted at an interface will depend on the colour (wavelength) of the light, and light of different colours incident at the same point and at the same angle will be refracted by different amounts. This effect is called dispersion.


Figure 8 The variation of the refractive index with wavelength for 'fused quartz' glass. (The wavelength is measured in a vacuum).

The dispersion of sunlight (which is a specific mix of visible wavelengths) in airborne water droplets is the fundamental cause of rainbows. A similar effect is shown in Figure 9, where a ray of white light (i.e. a similar mix of wavelengths to daylight) is shown splitting into a continuous band of colours as it passes through a glass prism. 10 等


Figure 9 A ray of white light is split into its constituent colours by dispersion in a prism.


Figure 8 The variation of the refractive index with wavelength for 'fused quartz' glass. (The wavelength is measured in a vacuum).

When Isaac Newton discovered this phenomenon he correctly interpreted it as indicating that white light was composed of all the colours that are usually seen in a rainbow. The extreme red and violet rays are shown and in between these are all the other colours that comprise the visible spectrum; orange, yellow, green, blue and indigo. Since the refractive index of glass decreases as the wavelength increases (as in Figure 8), red light will be refracted less at each interface than the shorter wavelength violet light, and will emerge travelling in a direction closer to that of the incident ray of white light.

## Question T4

A light ray is composed of a mixture of red and blue light, and is incident on a block of fused quartz glass at an angle of $40^{\circ}$. Use the data from Figure 8 to estimate the difference in the angles of refraction inside the quartz for the two colours. (Assume $\lambda=650 \mathrm{~nm}$ for red light and $\lambda=450 \mathrm{~nm}$ for blue light.)

## Question T5

A ray of light strikes the interface between air and a block of glass so that the angle of incidence is $35^{\circ}$. It is observed that the angle of refraction is $22^{\circ}$. Calculate the speed of light inside the glass block.


### 4.2 Image formation by a plane interface

If you look into a calm swimming pool you can see images of objects below the surface. Ray diagrams can be used to investigate these refracted images, just as they were used in the case of reflection. Figure 10 shows two rays from a point object O that are refracted at the interface between two different transparent materials. The two refracted rays appear to come from a virtual image at the point I'. Unfortunately, if we choose any other two rays from O , at different angles of incidence to the surface, we will find that they will appear to come from a virtual image which is slightly displaced from $I^{\prime}$. This means that the virtual image of a point object formed by refraction at a plane interface is never perfectly sharp; it suffers from what are called aberrations.


Figure 10 Formation of a virtual image by refraction at an interface.

Fortunately these aberrations are reduced when the eye receives refracted rays from only a small range of angles, and also when the angle of incidence is small. This is indicated in Figure 11, which shows an object at O in a material with refractive index $\mu_{2}$, viewed from directly above by an eye in a material with refractive index $\mu_{1}$. As drawn, the rays are refracted away from the normal as they enter material 1 , indicating that $\mu_{2}$ is greater than $\mu_{1}$. Using Snell's law it is possible to show that in this or similar situations, the real depth of the object below the interface (the distance ON ) is related to its apparent depth (the depth of the image I'N), by

$$
\frac{\text { real depth }}{\text { apparent depth }}=\frac{\mathrm{ON}}{\mathrm{I}^{\prime} \mathrm{N}}=\frac{\mu_{2}}{\mu_{1}}
$$

provided the angles of incidence and refraction are small.

Figure 11 The real and apparent depth of an object in a transparent material.


## Question T6

Derive Equation 3

$$
\frac{\text { real depth }}{\text { apparent depth }}=\frac{\mathrm{ON}}{\mathrm{I}^{\prime} \mathrm{N}}=\frac{\mu_{2}}{\mu_{1}}
$$

from Snell's law $\frac{\sin \theta_{1}}{\sin \theta_{2}}=\frac{\mu_{2}}{\mu_{1}} \quad$ (Eqn 1b)
An object O at the bottom of a swimming pool, viewed from above, appears to be 1 m below the surface. Use the data in Table 1 to find the

Table 1 The refractive indices of some transparent materials.

| Material | Refractive index |
| :---: | :---: |
| vacuum | 1.00 (exactly) |
| air* | 1.00029 |
| water | 1.33 |
| crown glass | 1.52 |
| fused quartz | 1.46 |
| diamond | 2.42 | real depth. (Hint: If $\theta$ is small, $\sin \theta \approx \tan \theta$.)



Analysis of the images of extended objects, especially when viewed at large angles of incidence, is more complicated in the case of refraction than reflection. We will not pursue the formation of such images systematically in this module but one of the complicating factors that may arise, total internal reflection, is dealt with in the next subsection.

### 4.3 Total internal reflection and optical fibres

When light of a particular wavelength passes from one transparent material to another of lower optical density (i.e. $\mu_{1}>\mu_{2}$ ), as shown in Figure 12, we are usually able to observe both reflection and refraction. From Snell's law we know that for any given ray the angle of refraction $\theta_{2}$ will generally be greater than the angle of incidence $\theta_{1}$ since $\mu_{1}>\mu_{2}$. In such cases Equation 1b
Snell's law $\frac{\sin \theta_{1}}{\sin \theta_{2}}=\frac{\mu_{2}}{\mu_{1}}$
can be rearranged to give:

$$
\begin{equation*}
\sin \theta_{1}=\frac{\mu_{2} \sin \theta_{2}}{\mu_{1}} \tag{4}
\end{equation*}
$$

The largest possible value of $\sin \theta_{2}$ is 1 and this occurs when $\theta_{2}=90^{\circ}$, that is when the refracted ray is parallel to the interface. Now, it is clear from Equation 4 that the corresponding value of $\sin \theta_{1}$ is $\mu_{2} / \mu_{1}$, which is less than 1 , so it follows that there is a limited range of angles of incidence that can give rise to refracted rays in medium 2.


Figure 12 Light rays of a particular wavelength and a range of angles of incidence pass from one material to another material of lower optical density.

The upper limit to this range, the angle of incidence which gives the refracted ray along the interface, is called the critical angle $\theta_{\mathrm{c}}$. Since the critical angle is the value of $\theta_{1}$ that ensures $\theta_{2}=90^{\circ}$ it follows from Equation 4

$$
\begin{equation*}
\sin \theta_{1}=\frac{\mu_{2} \sin \theta_{2}}{\mu_{1}} \tag{Eqn4}
\end{equation*}
$$

that

$$
\begin{equation*}
\theta_{\mathrm{c}}=\arcsin \left(\frac{\mu_{2}}{\mu_{1}}\right) \tag{5}
\end{equation*}
$$

The question that now arises is what happens to an incident ray which reaches the interface at an angle of incidence greater than $\theta_{c}$ ? According to Equation 4 there is no value of $\theta_{2}$ which can satisfy the equation (it would require $\sin \theta_{2}$ to be greater than 1) and so there is no refracted ray - the light cannot pass through into the material of lower optical density and so all the incident energy must be reflected at the boundary. We have here a mechanism for perfect $100 \%$ reflection! The phenomenon is called total internal reflection and it has many important applications.

## Question T7

A light ray passes from a block of glass $\left(\mu_{\text {glass }}=1.50\right)$ into water $\left(\mu_{\text {water }}=1.33\right)$. Find the critical angle $\theta_{c}$. Repeat the calculation for a ray passing from glass into air ( $\mu_{\mathrm{air}}=1.00$ ).

(a)


Figure 13 Total internal reflection in $45^{\circ}$ prisms. (a) A single prism, (b) two prisms separated as in a periscope, (c) two prisms as in prismatic binoculars.

Total internal reflection is exploited in many commercial applications where energy losses at reflecting surfaces must be minimized. For example, in Figure 13a we show a $45^{\circ}$ glass prism used as a prismatic reflector, in Figure 13b two such prisms as used in a periscope, and in Figure 13c two such prisms with their $45^{\circ}$ faces rotated through $90^{\circ}$, as might be found in prismatic binoculars.


Figure 13 Total internal reflection in $45^{\circ}$ prisms. (a) A single prism, (b) two prisms separated as in a periscope, (c) two prisms as in prismatic binoculars.

The prisms are made of glass with refractive index close to 1.5 so that their critical angle with air is $41.8^{\circ}$. When a ray enters such a prism, as shown, its angle of incidence within the glass is $45^{\circ}$ which is greater than the critical angle, so it is totally internally reflected and turned through $90^{\circ}$ at each reflection as shown. The energy loss at the air-glass interfaces is lower than that at a silvered mirror surface, so the images seen are brighter.

Perhaps the most important application of total internal reflection lies in the field of optical fibre technology. This is of vital importance in modern telecommunications networks. Optical fibres are usually made from very thin strands of quartz glass (often 0.2 mm or less in diameter), specially made to be as free as possible from impurities and imperfections. If a parallel beam of light (usually from a laser) enters one end of the fibre


Figure 14 A section of optical fibre and a confined light ray. at a small angle to the axis then the constituent light rays may be arranged to have large incidence angles (larger than the critical angle) at the fibre wall. Under these circumstances total internal reflections will prevent the light from escaping through the sides of the fibre and it can be conveyed over distances of the order of tens of kilometres to emerge at the other end without much loss of energy. A section of optical fibre, with an incident ray at a small angle to the axis of the fibre, is shown in Figure 14. The ray is confined within the fibre by successive reflections at the surface, provided the angle of incidence remains greater than the critical angle.

Optical fibres can carry an incredible amount of information. It is not the purpose of this module to describe exactly how this encoding is done but suffice it to say that the light often consists of very short pulses from a laser and these pulses can convey coded information in a variety of ways. In this application the optical fibre link behaves in many ways like the radiowave and microwave links that have been used for decades by the communications industry. Apart from the need for a fibre pathway between the source and the detector, the main difference is the much shorter wavelength (i.e. higher frequency) of the optical signal.
FLAP P6.2 Rays and geometrical optics

COPYRIGHT © $1998 \quad$ THE OPEN UNIVERSITY $\quad$ S570 V1.

This has the consequence that fibre-based systems can carry much more information than either radiowave or microwave links or even a normal wire connection.

For all their virtues, simple optical fibres may lose energy if their surfaces are damaged in any way. An improvement in the performance of optical fibres can therefore be made by preventing the light rays from ever touching the outer surface. This improved design can be achieved in two ways. First, a fibre can be constructed so that the refractive index decreases by a finite step at a fixed distance from the axis; the rays may then be totally reflected from this internal protected interface. This is called a stepped-index fibre and is illustrated in Figure 15.


Figure 15 A stepped-index fibre in which the refractive index changes at a fixed distance from the fibre axis (as indicated by the dashed line). Alternatively, a fibre can be made so that the refractive index decreases gradually with distance from the axis; this is called a graded-index fibre and will be discussed in the next subsection. The only major source of energy loss in the stepped-index or graded-index fibre is the absorption in the glass itself, which can be made very low.
Over the next few years optical fibres are likely to replace the hard-wire telephone-grid land lines throughout the UK and since 1988 a fibre optic undersea cable has been used for transatlantic telecommunication. The next phase of the transatlantic cable development (planned for 1996) will link the UK and France with two sites in the USA in a ring network involving two transatlantic fibre optic cables. Each of these cables will have two pairs of fibres, each $5 \mu \mathrm{~m}$ (or about ten wavelengths of light) in diameter.
FLAP P6.2 Rays and geometrical optics
$\begin{array}{llll}\text { COPYRIGHT © } 1998 & \text { THE OPEN UNIVERSITY } & \text { S570 V1.1 }\end{array}$

In each cable only one fibre pair is active, with the other pair as a spare in the event of damage. The power of this technology can be appreciated by realizing that, in principle, a single active pair of these fibres can carry the equivalent of 300000 simultaneous telephone conversations!

Sometimes many hundreds or thousands of optical fibres are packaged together to form a flexible fibre bundle that will confine rays within individual fibres, provided they are not bent into too tight a curve. These bundles may be used in large communication networks, where there are many sites for sending and receiving information. Fibre bundles can also be used to convey pictures in a very straightforward way. If a real image of an object is projected on to the end of a fibre bundle, each point in the image will correspond to a particular fibre and hence to a particular point at the output end of the bundle. So, if the relative fibre positions at each end of the bundle are the same, the image fed into one end of the bundle can be seen at the other end. These organized bundles are called coherent fibre bundles and they can be used, for example, in 'non-invasive' medicine for diagnosis, exploratory investigations or to assist surgeons.

### 4.4 Continuous refraction and atmospheric optics

To close this section it is interesting to consider what happens when a light ray travels through a medium in which the refractive index is changing continuously, rather than changing suddenly at an interface. Such situations lead to the phenomenon of continuous refraction which may arise, for example, when a ray of light travels through the atmosphere at an angle to the ground. The refractive index of air varies with density and temperature and, since both of these vary with altitude, the refractive index will also vary with altitude. Density is usually the dominant factor, as a result of which the refractive index usually decreases as the height above the ground increases.

Continuous refraction in the atmosphere has a number of interesting consequences. For example, when astronomers attempt to determine the precise position of a star they usually have to take the effect of continuous refraction into account. If the star is fairly close to the overhead point (the zenith) this is not too difficult; the curvature of the Earth can be ignored and the atmosphere can be treated as though it consisted of several horizontal layers with slightly different refractive indices.

The effect of this, as shown in Figure 16, is to make the star appear closer to the overhead point than it truly is.


Figure 16 Atmospheric refraction in a plane stratified atmosphere where the refractive index decreases as the altitude increases. The incident ray bends towards the optically denser air. (The effect has been greatly exaggerated for clarity.)
ground level
apparent direction
of star
air layers of increasing
refractive index

On the other hand the effect of continuous refraction on the light from stars near the horizon is much greater and much harder to calculate, partly because of the need to take account of the Earth's curvature but mainly because a greater part of the ray is in the lower part of the Earth's atmosphere where turbulence, localized atmospheric disturbances and pollution are all more likely to distort the light rays. This is part of the reason why sunsets are so dramatic and have so many colourful effects associated with them.

Figure 16 Atmospheric refraction in a plane stratified atmosphere where the refractive index decreases as the altitude increases. The incident ray bends towards the optically denser air. (The effect has been greatly exaggerated for clarity.)

$\leftrightarrow$ Does the atmosphere affect the time at which the Sun appears to set as seen from a fixed point on Earth? If so, would the Sun set earlier or later in the absence of atmospheric refraction?

Another atmospheric phenomenon associated with continuous refraction is the mirage familiar from films and advertisements where a weary desert traveller sees the illusion of water (or some similar liquid) in the distance. There are in fact several different kinds of mirage. The simplest may sometimes be seen when driving on a long straight road on a hot clear sunny day. You may notice what appear to be 'wet' patches on the road surface well ahead, but as you approach them they disappear and reappear further ahead. What you are seeing is a 'reflection' of the sky or some other object caused by continuous refraction in the layers of air near the hot road surface.

The ray diagram explaining such mirages is shown schematically in Figure 17; in this case the air closest to the ground is hotter and optically less dense than that above, so the usual arrangement is reversed and rays bend away from the ground. If the rate at which the temperature changes with height (i.e. the temperature gradient $d T / d h$ ) itself decreases appropriately with height then those rays that get closest to the ground will undergo the greatest degree of bending, as shown. The main effect takes place very near the road surface, where the temperature gradient in the air is quite large, but the angles of refraction are very small and so this mirage is only possible near to grazing incidence. That is why the 'wet' patches are usually seen far away on a flat road, though they can sometimes be seen much closer when you approach the brow of a hill. Convection currents in the air near the road surface cause the shimmering which usually accompanies this mirage effect.


Figure 17 A 'hot road' or 'desert' mirage.

## Question T8

A ray of light that rises into the Earth's atmosphere leaves the ground at approximately $45^{\circ}$ to the surface. If the refractive index of the atmosphere decreases gradually with height, how will refraction affect the path of the ray? How might this behaviour lead to a mirage?

It is possible to demonstrate continuous refraction in a simple laboratory experiment using a laser beam and a tank of sugar solution. The refractive index of a sugar solution depends on the concentration of the sugar. If a tank of sugar solution is created by allowing water to settle over a large amount of sugar and the system is then allowed to reach equilibrium, the sugar concentration will be highest immediately above the undissolved sugar and will fall smoothly with height above this. This creates a situation where the refractive index of the solution decreases with height. If a laser beam is shone into the tank the curved path of the continuously refracted beam through the solution can be easily seen and the mirage effect demonstrated.

Continuous refraction is also used in fibre optics. A fibre can be made so that the refractive index decreases gradually with distance from the axis; this is called a graded-index fibre. Figure 18 shows how a graded index fibre works. Rays at small angles to the axis always bend away from the optically less dense material near to the surface so they can progress down the fibre without ever touching the surface. $\stackrel{\text { diz }}{ }$ This has a similar power loss to the stepped-index fibre and in some applications is a preferred solution.


Figure 18 A graded-index optical fibre, with a refractive index that decreases as the distance from the axis increases.

## Question T9

What would happen to a light ray passing through a medium in which the refractive index changes continuously along the ray, but does so in such a way that at no point are there any changes in refractive index perpendicular to the ray?

## 5 Fermat's principle

Hero of Alexandria, who probably lived some time around 50 ( $\pm 30$ ) AD, proposed the following rule which he believed governed the propagation of light:

When light travels between two fixed points it does so in such a way that its path has the shortest possible length.

If correct, this rule can certainly explain rectilinear propagation; it would simply be a consequence of the geometrical fact that given any two points A and B (see Figure 19), a straight line provides the shortest path between them.


Figure 19 A light ray follows the shortest route between two fixed points. Any nearby path is longer.

Figure 20 shows how Hero's rule can be used to deduce the law of reflection. In this case the problem is to find the shortest path from $P$ and $\mathrm{P}^{\prime}$ that includes a reflection in the mirror. It is pretty clear that the solution must involve two straight lines, such as PA and $\mathrm{AP}^{\prime}$, but where must A be located to ensure that the total path length is as small as possible? As Hero realized (and as we will prove in Subsection 5.1), A must be located in such a way that $\theta_{\mathrm{i}}=\theta_{\mathrm{R}}$. Thus the particular path that obeys the law of reflection is also the path of shortest possible length, and the law of reflection may be viewed as a consequence of Hero's rule.


Figure 20 A conceivable route for light rays passing from P to $\mathrm{P}^{\prime}$ via a reflection at A ; but is it the route the rays will actually follow?

It took about 1600 years to take the next step. In 1657 the French mathematician Pierre de Fermat (1601-1665) published his now famous 'principle' which extended Hero's rule to include cases of refraction. If you look at Figure 21 it is obvious that the ray path PAP' between P and $\mathrm{P}^{\prime}$ is not the shortest path - that would be the straight line path from P to $\mathrm{P}^{\prime}$ and the ray does not follow that path. Now, although Hero's rule does not work in this case, Fermat realized that a simple modification to the ancient rule would lead to the known laws of reflection and refraction. According to Fermat the path taken by the light was not that of shortest length but rather that which required the least time. In the case of reflection, where only one medium is involved, the path of minimum time is also the path of minimum length. But in the case of refraction, where the two parts of the path are traversed in different media and therefore at different speeds, the minimum time condition can be used to predict the correct path but the minimum length condition cannot.

- If the light in Figure 21 followed a straight line path from P to $\mathrm{P}^{\prime}$, rather than the true path $\mathrm{PAP}^{\prime}$, would the time it spent in the medium with refractive index $\mu_{1}$ be increased? What can you say about the time spent in the medium of refractive index $\mu_{2}$ ?



Figure 21 The refraction of a light ray at a plane surface - Fermat's principle of least time can be used to determine the correct location of point A .

In its original formulation, Fermat's principle was stated as follows:
The actual path taken by a light ray in travelling between two points is the one that is traversed in the least time.

Fermat's principle presents an intriguing new perspective on the laws of geometrical optics, though it needs some further modification to enable it to deal with all the cases that arise in practice. For instance, it is now known that there are many situations involving highly curved reflectors in which the rays are such that the time taken to traverse them is greater than the time taken to traverse other conceivable rays. The somewhat technical reformulation of Fermat's principle that covers these and other such cases will now be stated and then explained.

The modern form of Fermat's principle is stated as follows:
If a light ray passes from one fixed point to another fixed point then the time taken to traverse the actual path will, to a first approximation, be equal to the time taken for light rays to traverse adjacent paths, i.e. the time taken to traverse the path will be stationary with respect to small variations in the path. ${ }^{\text {ang }}$

The key to understanding this form of Fermat's principle is the word stationary.

| FLAP | P6.2 | Rays and geometrical optics |  |
| :--- | ---: | :--- | ---: | :--- |
| COPYRIGHT | © 1998 | THE OPEN UNIVERSITY | S570 V1.1 |

Mathematically speaking a function $f(x)$ is said to be stationary at any point where the tangent to its graph is horizontal, and the point at which this occurs is called a stationary point.

(a)

(b)

Figure 22 The three situations in which a function may be stationary: (a) a local maximum, (b) a local minimum, and (c) a point of inflection with a horizontal tangent.

At such a point the gradient of the graph must be zero, i.e. the derivative $d f / d x=0$. The three situations in which functions become stationary are shown in Figure 22a, b and c; they are, respectively, referred to as a local maximum, a local minimum and a point of inflection (with a horizontal tangent). $\underline{\underline{2} 8 \mathrm{~g} 8}$ Since the tangent to the graph of $f(x)$ is horizontal in each of these cases it is clear that the function is neither increasing nor decreasing at the stationary point and that the values of $f(x)$ immediately to the right and to the left of the stationary point will be very close to the value of $f(x)$ at the stationary point itself (i.e. the value of $f(x)$ at the stationary point will be equal to the value of $f(x)$ at adjacent points, to a first approximation).

In the context of reflection and refraction (see Figures 20 and 21), the time $T$ taken to traverse a particular path from P to $\mathrm{P}^{\prime}$ via A will depend on the exact location of A, and this is determined by the value of $x$. Thus, we may write the time required to traverse any particular path as $T(x)$ and we may interpret Fermat's principle as asserting that the actual path followed by the light will be that which makes $T(x)$ stationary. 1098 In other words we can locate A and identify the actual path in either case by finding the value of $x$ for which $d T / d x=0$. This is exactly what we will do in the next two subsections.


Figure 20 A conceivable route for light rays passing from P to $\mathrm{P}^{\prime}$ via a reflection at A ; but is it the route the rays will actually follow?


Figure 21 The refraction of a light ray at a plane surface - Fermat's principle of least time can be used to determine the correct location of point A.

Study comment Subsections 5.1 and 5.2 contain proofs of the assertions made above but they assume some knowledge of differential calculus. If you are unfamiliar with calculus you may omit these two subsections without endangering your ability to meet the achievements of this module, provided you can state Fermat's principle and relate it to the laws of geometrical optics.

### 5.1 Reflection at a plane mirror

Consider the path of a light ray from P reflected in a plane mirror so that it passes through the point $\mathrm{P}^{\prime}$ as shown in Figure 20. The points P and $\mathrm{P}^{\prime}$ are at fixed distances $a$ and $b$ from the mirror and $d$ is the fixed length $\mathrm{NN}^{\prime}$. The light ray strikes the mirror at a point A and we let the distance NA be represented by the variable $x$, which defines the path of the ray. We assume that $\mathrm{P}, \mathrm{A}$ and $\mathrm{P}^{\prime}$ all lie in the same plane, although we could use Fermat's principle to prove this. Fermat's principle requires that the actual path has the particular value of $x$ such that $d T / d x=0$. We will show that this condition leads to the result $\theta_{\mathrm{i}}=\theta_{\mathrm{R}}$, as expected.


Figure 20 A conceivable route for light rays passing from P to $\mathrm{P}^{\prime}$ via a reflection at A ; but is it the route the rays will actually follow?

We will suppose that all the rays in Figure 20 are in a vacuum (though we could equally well deal with the case where they were all in some other uniform medium), so the time taken for the light to travel the distance $\left(\mathrm{PA}+\mathrm{AP}^{\prime}\right)$ is given by $T=\left(\mathrm{PA}+\mathrm{AP}^{\prime}\right) / c$.
FLAP P6.2 Rays and geometrical optics
$\begin{array}{llll}\text { COPYRIGHT © } 1998 & \text { THE OPEN UNIVERSITY } & \text { S570 V1.1 }\end{array}$

We can write the path $\left(\mathrm{PA}+\mathrm{AP}^{\prime}\right)$ in terms of $x$, with

$$
c T=\left(\mathrm{PA}+\mathrm{AP}^{\prime}\right)=\sqrt{a^{2}+x^{2}}+\sqrt{b^{2}+(d-x)^{2}}
$$

Differentiating this expression with respect to $x$ we have

$$
\begin{aligned}
& c \frac{d T}{d x}=\frac{1}{2}\left[2 x\left(a^{2}+x^{2}\right)^{-1 / 2}\right]-\frac{1}{2}\left\{2(d-x)\left[b^{2}+(d-x)^{2}\right]^{-1 / 2}\right\} \\
& c \frac{d T}{d x}=\frac{x}{\mathrm{PA}}-\frac{d-x}{\mathrm{AP}^{\prime}}
\end{aligned}
$$

If we set $d T / d x=0$, as is required by Fermat's principle, then

$$
\begin{equation*}
\frac{x}{\mathrm{PA}}-\frac{d-x}{\mathrm{AP}^{\prime}}=0 \tag{6}
\end{equation*}
$$

Since the angle $\mathrm{NP} A=\theta_{\mathrm{i}}$ and the angle $\mathrm{AP}^{\prime} \mathrm{N}^{\prime}=\theta_{\mathrm{R}}$ we can write

$$
\frac{x}{\mathrm{PA}}=\sin \theta_{\mathrm{i}} \quad \text { and } \quad \frac{d-x}{\mathrm{AP}^{\prime}}=\sin \theta_{\mathrm{R}}
$$

so Equation 6 gives $\sin \theta_{i}-\sin \theta_{R}=0$ i.e. $\theta_{i}=\theta_{R}$

Thus, as claimed earlier, the law of reflection is a consequence of Fermat's principle. 10 Ang allowable path with $\theta_{\mathrm{i}} \neq \theta_{\mathrm{R}}$ would take a greater time to traverse than the actual path. Though we have proved this mathematically you may still like to confirm it 'experimentally' - the next question will allow you to do so.

## Question T10

Consider a special case of reflection, as in Figure 20, with the distances: $a=10.0 \mathrm{~cm}, b=7.0 \mathrm{~cm}$ and $d=20.0 \mathrm{~cm}$. Use the laws of reflection to find the distance NA for a true reflection. Now, make a full scale drawing and, by directly measuring paths, check that this path length is less than that where the point A lies 2 cm nearer to N or 2 cm further from N, as predicted by Fermat's principle.


Figure 20 A conceivable route for light rays passing from P to $\mathrm{P}^{\prime}$ via a reflection at A ; but is it the route the rays will actually follow?

### 5.2 Refraction at a plane interface

We now give a similar derivation of the law of refraction. Remember, in Figure 21, the point P is in a transparent material of refractive index $\mu_{1}$ and the point $\mathrm{P}^{\prime}$ is in a material of refractive index $\mu_{2}$. The interface between the two media is a plane, and it is perpendicular to the plane containing the points $\mathrm{P}, \mathrm{P}^{\prime}, \mathrm{N}$ and $\mathrm{N}^{\prime}$. We will let $\mathrm{PN}=a, \mathrm{P}^{\prime} \mathrm{N}^{\prime}=b$ and $\mathrm{NN}^{\prime}=d$. The path of the light ray from P to $\mathrm{P}^{\prime}$ crosses the interface at point A, which is an unknown distance $x$ from N. Fermat's principle, applied to the path $\mathrm{PAP}^{\prime}$, should tell us the true value of $x$ and hence the relation between the incident and refracted angles $\theta_{1}$ and $\theta_{2}$. Again we assume that $\mathrm{P}, \mathrm{A}$ and $\mathrm{P}^{\prime}$ lie in a plane, but this could also be deduced from Fermat's principle. Remembering that the speed of light in a material is $c / \mu$, the time taken for the ray path $\mathrm{PAP}^{\prime}$ is given by:

$$
T=\mu_{1} \mathrm{PA} / c+\mu_{2} \mathrm{AP}^{\prime} / c
$$

so that $\quad c T=\mu_{1}\left(a^{2}+x^{2}\right)^{1 / 2}+\mu_{2}\left[b^{2}+(d-x)^{2}\right]^{1 / 2}$


Figure 21 The refraction of a light ray at a plane surface - Fermat's principle of least time can be used to determine the correct location of point A .

Differentiating this

$$
c T=\mu_{1}\left(a^{2}+x^{2}\right)^{1 / 2}+\mu_{2}\left[b^{2}+(d-x)^{2}\right]^{1 / 2}
$$

with respect to $x$ we find:

$$
c \frac{d T}{d x}=\mu_{1} \frac{1}{2}\left[2 x\left(a^{2}+x^{2}\right)^{-1 / 2}\right]-\mu_{2} \frac{1}{2}\left\{2(d-x)\left[b^{2}+(d-x)^{2}\right]^{-1 / 2}\right\}
$$

i.e. $\quad c \frac{d T}{d x}=\mu_{1} \frac{x}{\mathrm{PA}}-\mu_{2} \frac{d-x}{\mathrm{AP}^{\prime}}$

From Fermat's principle, we set $d T / d x=0$ and thus obtain

$$
\begin{equation*}
\mu_{1} \frac{x}{\mathrm{PA}}-\mu_{2} \frac{d-x}{\mathrm{AP}^{\prime}}=0 \tag{7}
\end{equation*}
$$

Also, since $\quad \frac{x}{\mathrm{PA}}=\sin \theta_{1} \quad$ and $\quad \frac{d-x}{\mathrm{AP}^{\prime}}=\sin \theta_{2}$
we have $\quad \mu_{1} \sin \theta_{1}=\mu_{2} \sin \theta_{2}$

Thus, as claimed earlier, the law of refraction is a consequence of Fermat's principle. 1 principle has given us Snell's law of refraction. Any closely similar path from P to $\mathrm{P}^{\prime}$ would take a longer time than the actual path.
We now refer back to Figure 4, which shows the formation of a real image by reflection at a plane mirror followed by refraction in a human eye. In this case there are many different rays from a point on the object to the corresponding point on the (real) image. Since all of these rays are equally valid no one of them can be singled out as corresponding to the minimum (or maximum) time; indeed, we can conclude from Fermat's principle that light takes the same amount of time to traverse each of these rays. This situation must generally occur when a real image is formed and provides another way of thinking about images and about Fermat's principle: all rays from a point on an object to the corresponding point on its image must take the same time to traverse.

Figure 4 A bundle of rays diverging from O and reflected by a plane mirror, forms a virtual image at $\mathrm{I}^{\prime}$ behind the mirror, while the human eye forms a real image at I.


### 5.3 The principle of reversibility

It was noted earlier, in Subsections 3.1 and 4.1, that the directions of rays can always be reversed. Though perhaps not as important as the rules of reflection and refraction, this too is a general principle of geometrical optics and is usually referred to as the principle of reversibility.

It may be stated formally as follows:
The result of reversing the direction of a light ray is another valid light ray.
Fermat's principle provides a general basis for the principle of reversibility. Since the time light requires to traverse any part of a path depends only on the length of the path and the refractive index along that path, it is quite clear that the direction in which the light is travelling plays no role in determining the travel time. Thus, if a particular path is the path of stationary travel time for light travelling in one direction, it will also be the path of stationary travel time for light travelling in the opposite direction. Thus all rays should be reversible, as the principle asserts.

### 5.4 The physical basis of Fermat's principle

So far we have stated Fermat's principle and shown that it is consistent with the principles of geometrical optics (the principles of rectilinear propagation and reversibility, and the laws of reflection and refraction). If we were to ask for a physical justification of Fermat's principle we would have to resort to a discussion of the (electromagnetic) wave model of light. Whilst a full discussion of this would be out of place in this module, it is interesting to point out some features of the argument, since it gives a valuable insight into the relationship between the full wave model of light and the ray approximation.

Imagine an idealized ray of light as a line joining two fixed points P and $\mathrm{P}^{\prime}$. This path is a straight line in any material with constant refractive index. Imagine also a second route between P and $\mathrm{P}^{\prime}$ which deviates from the first only by very small amounts. Fermat's principle requires that, to a first approximation, the second route has the same journey time as the first since the first corresponds to a stationary journey time. From the perspective of the wave model, it follows that waves that start out in phase from P and travel along the two different routes will, to a first approximation, still be in phase when they arrive at $\mathrm{P}^{\prime}$. Such waves will tend to reinforce one another when they arrive, giving constructive interference.

This implies that there will be a group of wavelets travelling close to the actual path which will add constructively and maximum energy can be transmitted along the path. In contrast, if we consider any other path from P to $\mathrm{P}^{\prime}$ which does not correspond to a stationary journey time then slightly different nearby paths are less likely to have similar journey times and it is unlikely that waves starting out in phase from P and travelling along those paths will interfere constructively when they arrive at $\mathrm{P}^{\prime}$. Thus, the wave phenomenon of interference provides a justification for Fermat's principle. What is required for the actual path of the ray is that nearby rays have the same journey time, not necessarily that this time should be a minimum - it could equally well be a maximum or any stationary value. The condition that nearby paths have the same journey time, to first approximation, is simply that $d T / d x=0$.

## 6 Closing items

### 6.1 Module summary

1 Although light is fundamentally a wave phenomenon there are important simplifications which can be made when the wavefronts are not restricted by apertures or obstacles whose size is comparable with the wavelength. Under these conditions diffraction is usually negligible and the ray approximation is valid. This is the domain of geometrical optics.
2 A ray is a directed line (i.e. a line with an arrow on it) showing a possible path along which light can travel and indicating the direction in which such light would transport energy. According to the principle of rectilinear propagation light travels in straight lines in a uniform medium. According to the principle of reversibility the result of reversing the direction of a light ray is another valid light ray.
3 According to the law of reflection when a light ray is reflected from a surface, (i) the incident ray, the reflected ray and the normal all lie in the same plane, and (ii) the angle of incidence is equal to the angle of reflection: $\theta_{i}=\theta_{\mathrm{R}}$.
4 When light from an object is reflected in a plane mirror a reversed virtual image is formed. This image is erect, the same size as the original object and located as far behind the mirror as the object is in front.

5 When light travels through a transparent medium it always travels more slowly than in a vacuum and the refractive index $\mu$ of the medium is defined as the ratio of the speed of light in a vacuum to the speed of light in the medium, i.e. $\mu=c / v$.
6 According to the law of refraction when a light ray is refracted at the interface between two transparent media with different refractive indices, (i) the incident ray, the refracted ray and the normal all lie in the same plane, and (ii) the angle of incidence and the angle of refraction are related by Snell's law

$$
\frac{\sin \theta_{1}}{\sin \theta_{2}}=\frac{\mu_{2}}{\mu_{1}}
$$

## (Eqn 1b)

7 When a light ray reaches a boundary with a less optically dense medium (i.e. one of lower refractive index) there is a critical angle of incidence, above which total internal reflection occurs. Optical fibres are an important commercial application of total internal reflection.
8 When a light ray travels through a region where there is a smooth variation in the refractive index the ray is continuously refracted and may follow a curved path. This effect explains various optical effects of the atmosphere, such as mirages, and finds commercial applications in the production of graded-index optical fibres.

9 According to Fermat's principle, if a light ray passes from one fixed point to another fixed point then the time taken to traverse the actual path will, to a first approximation, be equal to the time taken for light rays to traverse adjacent paths, i.e. the time taken to traverse the path will be stationary with respect to small variations in the path. This principle can be justified in terms of the wave model of light and provides a basis for geometrical optics from which the principles of rectilinear propagation and reversibility, and the laws of reflection and refraction can all be deduced.

### 6.2 Achievements

Having completed this module, you should be able to:
A1 Define the terms that are emboldened and flagged in the margins of the module.
A2 Recall and use the law of reflection at plane mirrors, and describe the image formed at a plane mirror.
A3 Define the refractive index of a transparent material. Recall and use the law of refraction (including Snell's law), and describe the image formed at a plane interface between transparent media.
A4 Calculate the critical angle at appropriate interfaces, describe the phenomenon of total internal reflection, and explain how this phenomenon is used in fibre optics and in other optical systems.
A5 Describe the phenomenon of continuous refraction and explain how it is involved in the operation of graded-index optical fibres and in atmospheric phenomena such as mirages.
A6 State the modern form of Fermat's principle. Justify it in terms of the wave model of light, and explain (without proof) its relation to the principles of rectilinear propagation and reversibility, and to the laws of reflection and refraction.

Study comment You may now wish to take the Exit test for this module which tests these Achievements. If you prefer to study the module further before taking this test then return to the Module contents to review some of the topics.
$\begin{array}{llll}\text { COPYRIGHT © } 1998 & \text { THE OPEN UNIVERSITY } & \text { S570 V1.1 }\end{array}$

### 6.3 Exit test

Study comment Having completed this module, you should be able to answer the following questions each of which tests one or more of the Achievements.

## Question E1

(A2 and A3) State the principles of rectilinear propagation and reversibility, and the laws of reflection and refraction.

## Question E2

(A1) Explain what is meant by the terms virtual image and real image of a point object. Give one example of each kind of image.

## Question E3

(A2) When you look in a mirror your left hand appears to be your mirror image's right hand. Why is the mirror image not also upside down?

## Question E4

(A3) Draw a ray diagram to show how a person, looking directly down into a pool of water, can see a virtual image of an object at the bottom. If the real depth of the pool is 2.00 m what is the apparent depth of the object, given that the refractive index of water is 1.33 ?

## Question E5

(A3) A ray of light passes from a material with refractive index $\mu_{1}$ into another with refractive index $\mu_{2}$ and then into a third material with index $\mu_{3}$. Show that the ray finally emerges with the same angle as if it had passed directly from the first to the third material. (Assume that all the interfaces are parallel.) Explain whether this would be true for any number of parallel layers of different transparent materials.


Figure 13 Total internal reflection in $45^{\circ}$ prisms. (a) A single prism, (b) two prisms separated as in a periscope, (c) two prisms as in prismatic binoculars.

## Question E6

(A4) A light ray passes into a right-angled prism in the way shown in Figure 13a. Calculate the smallest value of the refractive index of the glass for which total internal reflection would take place when the prism is in air ( $\mu_{\text {air }}=1.00$ ).

## Question E7

(A3) A light ray in air strikes a glass block at an angle of $15.0^{\circ}$ to the normal. The ray is refracted into the block at an angle of $9.9^{\circ}$. Calculate the refractive index of the glass.

## Question E8

(A5) What is meant by saying that a light ray can be continuously refracted? Describe the conditions in which such a phenomenon might occur.


## Question E9

(A0) State Fermat's principle and explain (without proof) how it is related to the principles of rectilinear propagation and reversibility, and the laws of reflection and refraction.

Study comment This is the final Exit test question. When you have completed the Exit test go back to Subsection 1.2 and try the Fast track questions if you have not already done so.

If you have completed both the Fast track questions and the Exit test, then you have finished the module and may leave it here.

| FLAP | P6.2 | Rays and geometrical optics |  | S570 V1.1 |
| :--- | ---: | :--- | :--- | :--- |
| COPYRIGHT © 1998 | THE OPEN UNIVERSITY |  |  |  |

