UNIT 2 REFLECTION AND REFRACTION OF LIGHT

Structure

2.1 INTRODUCTION

In the previous unit you have learnt that light is an electromagnetic wave. It is made up of mutually supporting electric and magnetic fields, which vary continuously in space and time. An interesting question related to e.m. waves is: What happens to these fields when such a wave is incident on the boundary separating two optically different media? From Unit 7 of PHE-02 Course you may recall that when a wave passes from air to water or air to glass, we get a reflected wave and a refracted wave. Reflection of light from a silvered surface, a looking mirror say, is the most common optical effect. Reflection of e.m. waves governs the working of a radar. Reflection of radiowaves by the ionosphere makes signal transmission possible and is so crucial in the area of communication.

In your earlier school years you have learnt that refraction explains the working of lenses and is responsible for seeing; our contact with surroundings. Even the grand spectacle of sun-set or a rainbow can be explained in terms of refraction of light. Refraction of e. m. waves forms the basis of one of the greatest technological applications in signal transmission. In fact, electro-optics has seen tremendous growth via optical fibres for a variety of applications.

In Unit 7 of PHE-02 course on Oscillations and Waves, you learnt to explain reflection and refraction of waves on the basis of Huygens' wave model. Now the question arises: Can we extend this analysis to electromagnetic waves, which include visible light, radiowaves, microwaves and X-rays? In Sec. 2.2 you will learn to derive the equations for reflected and transmitted fields (E and B) when an e.m. wave is incident normally as well as obliquely on the boundary of two media.

You are aware that many physical systems behave according to optimisation principle. In PHE-06 course you have learnt that-when several fluids'at different temperatures are mixed, the heat exchange takes place so that the total entropy of the system is maximum. A ball rolling on an undulating surface comes to rest at the lowest point. The profoundness of such situations and scientific laws governing them led Fermat to speculate: Does light also obey some optimization principle? And he concluded: **Ray of light chooses a path of extremum between two points.** This is known as Fermat's principle, Implicit in it are the assumptions

(i) Light travels at a finite speed, and

(ii) The speed of light is lower in a denser medium.

In Sec. 2.4 you will learn about Fermat's principle. We have shown that all laws of geometrical optics are contained in it.

Objectives

After studying this unit you should be able to

- explain reflection and refraction of e.m. waves incident normally and obliquely on the interface separating two optically different media
- apply Fermat's principle to explain the reflection and refraction of light, and
- solve problems based on reflection and refraction of e.m. waves.

2.2 ELECTROMAGNETIC WAVES AT THE INTERFACE SEPARATING TWO MEDIA

Consider a plane electromagnetic wave that is incident on a boundary between two linear media. That is, D and H are proportional to E and B, respectively, and the constants of proportionality are independent of position and direction. You can visualise it as light passing from air (medium 1) to glass (medium 2). Let us assume that there are no free charges or currents in the materials.

Fig. 2.1 shows a plane houndary between two media having different permittivity

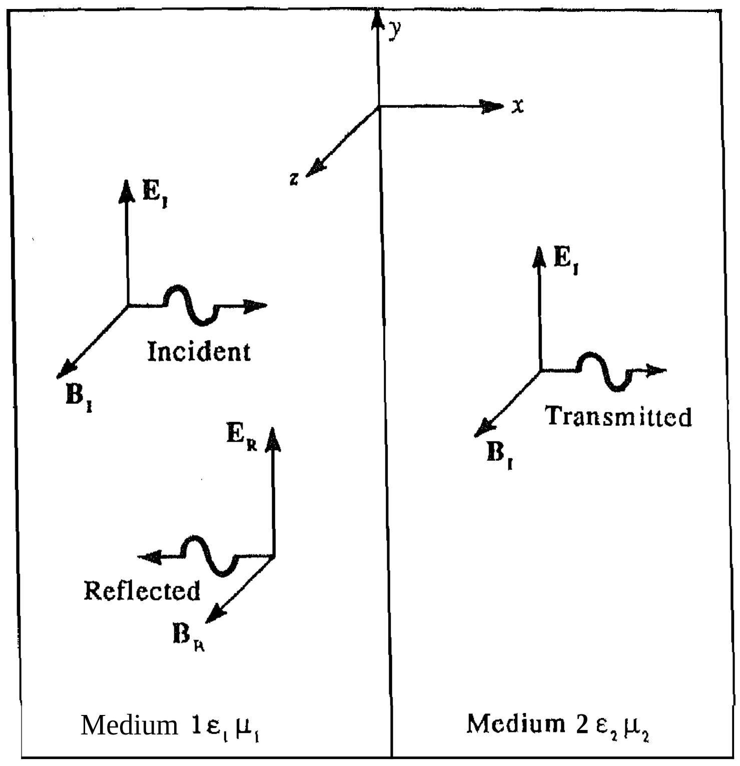


Fig.2.1: A uniform plane wave is incident normally on a plane boundary. The reflected and refracted (transmitted) waves are also shown. The angle of incidence is α and angle of refraction is β.

and permeability: ε_1, μ_1 for medium 1 and ε_2, μ_2 for medium 2. A uniform plane wave travelling to the right in medium 1 is incident on the interface **normal** to **the** boundary. As in the case of waves on a string, we expect a reflected wave propagating back into the medium and a transmitted (or refracted) wave travelling in the second medium, We wish (i) to derive expressions for the fields associated with reflected and refracted waves in terms of the field associated with the incident wave and (ii) know the fraction of the incident energy that is reflected and transmitted. To do so we need to know the **boundary conditions** satisfied by these waves at the interface separating the two media. We obtain these conditions by stipulating that Maxwell's equalions must be satisfied at the boundary between these media. We first state the appropriate conditions. Their proof is given in the appendix to this Unit.

Boundary Conditions

You learnt to derive the boundary conditions from Maxwell's equations for a medium free of charges and currents in Unit 15 of the PHE-07 course on electric and magnetic phenomena. For your convenience, we rewrite appropriate integral form of these equations:

$$\varepsilon \int_S \mathbf{E} \cdot d\mathbf{S} = 0 \tag{2.1a}$$

$$\int_S \mathbf{B} \cdot d\mathbf{S} = 0 \tag{2.1b}$$

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = \frac{d}{dt}\int_s \mathbf{B} \cdot d\mathbf{S} \tag{2.1c}$$

and

$$\frac{1}{\mu} \oint_C \mathbf{B} \cdot d\mathbf{l} = \varepsilon \frac{d}{dt}\int_S \mathbf{E} \cdot d\mathbf{S} \tag{2.1d}$$

where S is a surface bound by the closed loop C.

The electric field can oscillate either parallel or normal to the plane of incidence. The magnetic field B will then be normal or parallel to the plane of incidence. We will denote these with subscripts $\parallel$ (parallel) and $\perp$ (normal). The boundary conditions for normal and parallel components of electric and magnetic fields take the form (Appendix **A).**

$$\varepsilon_1 E_{1\perp} - \varepsilon_2 E_{2\perp} = 0 \tag{2.2a}$$

$$B_{1\perp} - B_{2\perp} = 0 \tag{2.2b}$$

$$E_{1\parallel} - E_{2\parallel} = 0 \tag{2.2c}$$

and

$$\frac{1}{\mu_1} B_{1\parallel} - \frac{1}{\mu_2} B_{2\parallel} = 0 \tag{2.2d}$$

We shall now use the boundary conditions expressed by Eqs. (2.2a- d) to study reflection and refraction (transmission) at normal as well as oblique incidence.

2.2.1 Normal Incidence

Refer to Fig. 2.2. The yz-plane ($x = 0$) forms the interlace of two optically transparent (non-absorbing) media (refractive indices n_1 and n_2). **A** sinusoidal plane wave of frequency ω travelling in x-direction is incident from the left. From Unit 7 of the Oscillations and Waves course you will recall that progressive waves are partially reflected and partially refracted at the boundary separating two physically different media. However, the energy of the reflected or transmitted e.m. waves depends upon their refractive indices.

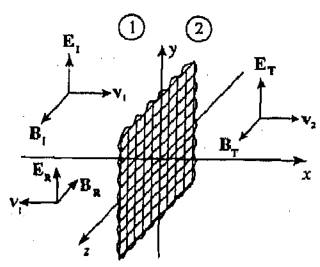


Fig.2.2: A sinusoidal plane e.m. wave Incident normally at the boundary of two optically transparent media

The appropriate magnetic fields to be associated with electric fields are obtained from the equation

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

Let us suppose that the electric field is along the y-direction. Then the electric and magnetic fields associated with the incident wave are given by

$$\mathbf{E_I}(x, t) = E_{0I}\hat{\mathbf{j}} \exp [i(k_I x - \omega t)] \qquad (2.3a)$$

and

$$\mathbf{B_I}(x, t) = \frac{E_{0I}}{v_1} \hat{k} \exp [i(k_I x - \omega t)] \qquad (2.3b)$$

The reflected wave propagates back into the first medium and can be represented by the following fields:

$$\mathbf{E_R}(x, t) = E_{oR}\hat{\mathbf{j}} \exp [-i(k_I x + \omega t)] \qquad (2.4a)$$

and

$$\mathbf{B_R}(x, t) = -\frac{E_{oR}}{v_1} \hat{k} \exp [-i(k_I x + \omega t)] \qquad (2.4b)$$

The minus sign in the exponents in Eqs. **(2.4a,b)** indicates that propagation of the wave is in the **-x direction.** But the negative sign with the amplitude in Eq. **(2.4b)** arises because of transverse nature of **e.m.** waves and that the electric and magnetic field vectors should obey the relation

$$\mathbf{B_R} = \frac{1}{v_1} (\hat{\mathbf{k}}_\mathbf{I} \times \mathbf{E_R})$$

where $\hat{\mathbf{k}}_\mathbf{I}$ is unit vector along the direction of incidence.

If you visualise Eqs. (2.3) and (2.4) diagramatically, you will note that the electric vectors have been kept fixed in the same direction but the magnetic field vectors have been oriented. The orientation of the magnetic field vector ensures that the flow of energy is always along the direction of propagation of the wave (Poynting theorem).

The electric and magnetic fields of the transmitted wave, which **travels** to the right in medium 2, are given by

$$\mathbf{E_T}(x, t) = E_{oT}\hat{\mathbf{j}} \exp [i(\omega_T t + k_T x)] \qquad (2.5a)$$

and

$$\mathbf{B_T}(x, t) = \frac{1}{v_2}\left[\hat{\mathbf{k}}_\mathbf{T} \times \mathbf{E_T}(x, t)\right] \qquad (2.5b)$$

The phenomenon of reflection and refraction is usually analysed in two parts:

(i) To determine the relations between the field vectors of the reflected and refracted waves in terms of that of the incident wave. These relations determine the reflection and the transmission coefficients. In this derivation, we match the E and **B** fields in the two media at the interface with the help of appropriate boundary conditions there.

(ii) To establish relations between the angle of incidence and the angles of reflection and refraction we may emphasize that so far as the laws of reflection and refraction are concerned, explicit use of any boundary condition is not required.

Fresnel's Amplitude Relations

To derive expressions for the amplitudes of the reflected and the refracted waves in terms of the amplitude of the incident wave, we apply boundary conditions given by Eq. (2.2a-d) at **every point** on the interface at **all times.** At $x = 0$, the combined field to the left ($\mathbf{E}_I + \mathbf{E}_R$ and $\mathbf{B}_I$ and $\mathbf{B}_R$) must join the fields to the right ($\mathbf{E}_T$ and $\mathbf{B}_T$). For normal incidence, there are no normal field components (perpendicular to the interface). But why? This is because neither E nor B field is in the x-direction. This means that Eqs. (2.2a,b) are trivial and only tangential components of the electric and magnetic fields should be matched at the plane $x = 0$. Thus

$$E_{0I} + E_{0R} = E_{0T} \tag{2.6a}$$

and

$$\frac{1}{\mu_1}(B_{0I} + B_{0R}) = \frac{1}{\mu_2} B_{0T}$$

or

$$\frac{1}{\mu_1}\left(\frac{E_{0I}}{v_1} - \frac{E_{0R}}{v_1}\right) = \frac{1}{\mu_2}\frac{E_{0I}}{v_2}$$

which, on simplification yields

$$E_{0I} - E_{0R} = \alpha E_{0T} \tag{2.6b}$$

where

$$\alpha = \frac{\mu_1 v_I}{\mu_2 v_2} = \sqrt{\frac{\mu_1 \varepsilon_1}{\mu_2 \varepsilon_2}} = \frac{\mu_1 n_2}{\mu_2 n_1} \tag{2.6c}$$

Solving Eqs. (2.6a) and (2.6b) for the reflected and transmitted electric field amplitudes in terms of the incident amplitude, you will find that

$$E_{0R} = \left(\frac{1-\alpha}{1+\alpha}\right) E_{0I} \tag{2.7a}$$

and

$$E_{0T} = \frac{2}{1+\alpha} E_{0I} \tag{2.7b}$$

For most optical media, the permeabilities are close to their values in vacuum ($\mu_1 \approx \mu_2 \approx \mu_0$). In such cases $a = \frac{v_1}{v_2}$ and we have

$$E_{0R} = \left(\frac{v_2 - v_1}{v_2 + v_1}\right) E_{0I}$$

and

$$E_{0T} = \frac{2 v_2}{v_2 + v_1} E_{0I} \tag{2.8}$$

This suggests that when $v_2 > v_1$, the reflected wave will be in phase with the incident wave and for $v_2 < v_1$, the reflected and incident waves will be out of phase. This is illustrated in Fig. 2.3.

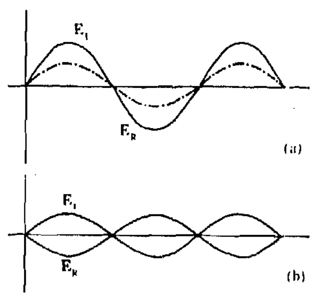


Fig.2.3: The phase relationship between reflected wave and the incident wave

In terms of the index of refraction $n \left(= \frac{c}{v} \right)$, we can rewrite Eq. (2.8) as

$$E_{0R} = \frac{n_1 - n_2}{n_1 + n_2} E_{0I}$$

and

$$E_{0T} = \frac{2n_1}{n_1 + n_2} E_{0I} \tag{2.9}$$

When an e.m. wave passes from a rarer medium to a denser medium ($n_1 < n_2$), the ratio $\frac{E_{0R}}{E_{0I}}$ will be negative. Physically, it means that the reflected wave is 180° out of phase with the incident wave. You have already learnt it in case of reflection of sound waves in the course on Oscillations and Waves. When an e.m. wave is incident from a denser medium on the interface separating it from a rarer medium ($n_1 > n_2$), the ratio $\frac{E_{0R}}{E_{0I}}$ is positive and no such phase change occurs.

We can now easily calculate the **reflection** and the **transmission coefficients**, which respectively measure the fraction of incident energy that is reflected and transmitted. The first step in this calculation is to recall that

$$R = \frac{I_R}{I_I}$$

and

$$T = \frac{I_T}{I_I}$$

where I_R, I_T and I_I respectively denote the reflected, transmitted and incident wave intensity. Intensity is defined as the average power per unit area, $(1/2)\, v\, E^2$. So you can readily show that

$$R = \frac{I_R}{I_I} = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2 \tag{2.10a}$$

and

$$T = \frac{I_T}{I_I} = \frac{n_2}{n_1} \left(\frac{2n_1}{n_1 + n_2} \right)^2 \tag{2.10b}$$

You can convince yourself that R $+T = 1$. For air ($n_1 = 1$) - glass ($n_2 = 1.5$) interface, the R and T coefficients have the values R = 0.04 and T = 0.96. There is no energy stored (or absorbed) at the interface and you can now realise why most of the light is transmitted.

We will now repeat this exercise for the case of oblique incidence.

2.2.2 Oblique Incidence

Refer to Fig. 2.4. A plane electromagnetic wave is incident at an angle θ_I. Let the angles of reflection and refraction be θ_R and θ_T. We can represent the fields associated with these three plane electromagnetic waves as

Incident Wave

$$\mathbf{E_I} = \mathbf{E_{0I}} \exp\left[-i\left(\omega_I t - \mathbf{k_I}\cdot\mathbf{r}\right)\right]$$

$$\mathbf{B_I} = \frac{1}{v_1}\left(\hat{\mathbf{k}}_\mathbf{I} \times \mathbf{E_I}\right) \tag{2.11a}$$

Reflected Wave

$$\mathbf{E_R} = \mathbf{E_{0R}} \exp\left[-i\left(\omega_R t - \mathbf{k_R}\cdot\mathbf{r}\right)\right]$$

$$\mathbf{B_R} = \frac{1}{v_1}\left(\hat{\mathbf{k}}_\mathbf{R} \times \mathbf{E_R}\right) \tag{2.11b}$$

Transmitted Wave

$$\mathbf{E_T} = \mathbf{E_{0T}} \exp\left[-i\left(\omega_T t - \mathbf{k_T}\cdot\mathbf{r}\right)\right]$$

$$\mathbf{B_T} = \frac{1}{v_2}\left(\hat{\mathbf{k}}_\mathbf{T} \times \mathbf{E_T}\right) \tag{2.11c}$$

You may recall that the boundary conditions must hold at every point on the interface at all times. If the boundary conditions hold at a point and at sometime, they will hold at all points in space for all subsequent times only if the exponential parts in above expressions for each wave are the same, i.e.

$$\omega_I t - \mathbf{k_I}\cdot\mathbf{r} = \omega_R t - \mathbf{k_R}\cdot\mathbf{r} = \omega_T t - \mathbf{k_T}\cdot\mathbf{r}$$

at the interface. This implies that: for equality of phases at all times we must have

$$\omega_I = \omega_R = \omega_T = \omega \quad (\text{say}) \tag{2.12a}$$

That is, the frequency of an e.m. wave does not change when it undergoes reflection and refraction: all waves have the same frequency. Since the fields must satisfy Maxwell's equations, we must have for the wave vectors

$$\frac{k_I^2}{\omega^2} = \frac{1}{c^2} = \varepsilon_1 \mu_1 \tag{2.13a}$$

$$\frac{k_T^2}{\omega^2} = \frac{1}{c^2} = \varepsilon_2 \mu_2 \tag{2.13b}$$

(2.13c)

Further, let k_{Ix}, k_{Iy} and k_{Iz} represent the x, y and z components of $\mathbf{k}_I$. We can use similar notation for $\mathbf{k}_T$ and $\mathbf{k}_R$. For the continuity conditions to be satisfied at all points on the interface, we must have

$$k_{Iy} = k_{Ty} = k_{Ry} \qquad (2.14a)$$

and

$$k_{Iz} = k_{Tz} = k_{Rz} \qquad (2.14b)$$

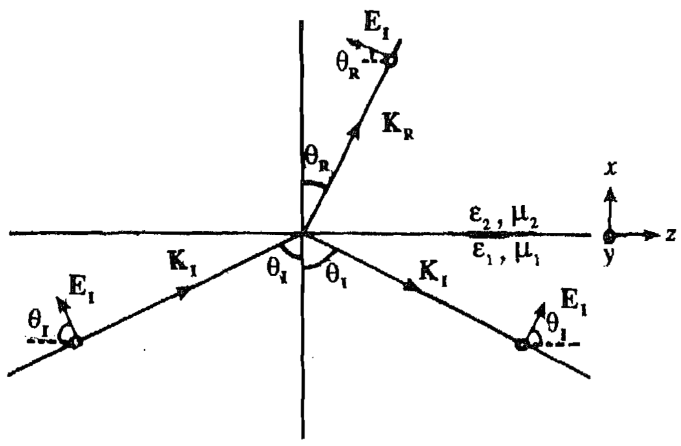


Fig.2.4: The reflection of a plane wove with its electric vector parallel to the plane of incidence

Let us choose the y-axis such that

$$k_{Iy} = 0$$

(i.e. we assume $\mathbf{k}_I$ to lie in the x-z plane - see Fig. 2.4). Consequenlly

$$k_{Ty} = k_{Ry} = 0 \qquad (2.14\sim)$$

This result implies that the vectors $\mathbf{k}_I$, $\mathbf{k}_T$ and $\mathbf{k}_R$ will lie in the same plane. Further, from Eq. (2.14b) we get

$$k_I \sin\theta_I = k_T \sin\theta_T = k_R \sin\theta_R \qquad (2.15)$$

Since $|\mathbf{k}_I| = |\mathbf{k}_R|$ (see Eq. 2.13a and c), we must have

$$\theta_I = \theta_R \qquad (2.16)$$

That is, the angle of incidence is equal to the angle of reflection, which is the law of reflection. Further,

$$\frac{\sin\theta_I}{\sin\theta_T} = \frac{k_T}{k_I} = \frac{\omega\sqrt{\varepsilon_2\mu_2}}{\omega\sqrt{\varepsilon_1\mu_1}}$$

or

$$\frac{\sin\theta_I}{\sin\theta_T} = \sqrt{\frac{\varepsilon_2\mu_2}{\varepsilon_1\mu_1}} \qquad (2.17)$$

If we denote the speeds of propagation of the waves in media 1 and 2 by $v_1\left(=\frac{1}{\sqrt{\varepsilon_1\mu_1}}\right)$ and $v_2\left(=\frac{1}{\sqrt{\varepsilon_2\mu_2}}\right)$ we find that Eq. (2.17) can be rewritten as

$$\frac{\sin\theta_I}{\sin\theta_T} = \frac{v_1}{v_2} = \frac{n_2}{n_1} \tag{2.18}$$

where $n_1 = \frac{c}{v_1} = c\sqrt{\varepsilon_1 \mu_1}$ and $n_2 = \frac{c}{v_2} = c\sqrt{\varepsilon_2 \mu_2}$.

represent the refractive indices of media 1 and 2 respectively. Do you recognise Eq. (2.18)? It is the well known **Snell's law.**

Eqs. (2.16) and (2.18) constitute the **laws of reflection and refraction** in optics.

You can now derive Fresnel's amplitude relations following the procedure outlined for the case of normal incidence. For brevity, we just quote the results without going into details. (You will not be examined for the same in the term-end examination.) When E oscillates parallel to the plane of incidence, we have

$$\frac{E_{R\parallel}}{E_{I\parallel}} = \frac{\tan(\theta_I - \theta_T)}{\tan(\theta_I + \theta_T)} \tag{2.19a}$$

$$\frac{E_{T\parallel}}{E_{I\parallel}} = \frac{2\cos\theta_I \sin\theta_T}{\sin(\theta_I + \theta_T)\cos(\theta_I - \theta_T)} \tag{2.19b}$$

When E oscillates normal to the plane of incidence, we have

$$\frac{E_{T\perp}}{E_{I\perp}} = \frac{\sin(\theta_I - \theta_T)}{\sin(\theta_I + \theta_T)} \tag{2.20a}$$

$$\frac{E_{T\perp}}{E_{I\perp}} = \frac{2\sin\theta_T \cos\theta_I}{\sin(\theta_I + \theta_T)} \tag{2.20b}$$

You can easily verify that for normal incidence these equations reduce to Eq. (2.9).

The corresponding expressions for reflections and transmission coefficients for normal and parallel oscillations of E when a plane wave is incident obliquely are

$$R_{\parallel} = \frac{\tan^2(\theta_I - \theta_T)}{\tan^2(\theta_I + \theta_T)} \tag{2.21a}$$

$$T_{\parallel} = \frac{\sin 2\theta_I \sin 2\theta_T}{\sin^2(\theta_I + \theta_T)\cos^2(\theta_I - \theta_T)} \tag{2.21b}$$

$$R_{\perp} = \frac{\sin^2(\theta_I - \theta_T)}{\sin^2(\theta_I + \theta_T)} \tag{2.21c}$$

and

$$T_{\perp} = \frac{\sin 2\theta_I \sin 2\theta_T}{\sin^2(\theta_I + \theta_T)} \tag{2.21d}$$

As before, you can easily show that for normal incidence these equations reduce to **Eq.** (2.10a, b).

2.3 IDEALIZATION OF WAVES AS LIGHT RAYS

So far you have learnt to explain reflection and refraction of plane electromagnetic waves at a plane interface. This signifies a relatively simple situation where the solutions of Maxwell's equations give the laws of propagation of light. It is not true in general and we invariably seek approximations to describe a phenomenon well. One such approximation makes use of smallness of wavelength of light. You know that the wavelength of light is very small ($\sim 10^{-7}$m). It is orders of magnitude less

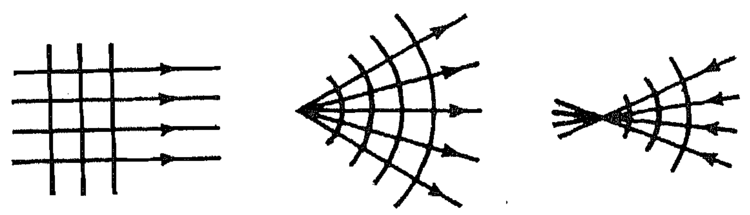


Fig.2.5: Ray representation of n plane, diverging spherical and converging spherical wavefronts moving from left to right

than the dimensions of optical instruments such as telescopes and microscopes. In such cases, the passage of light is most easily shown by geometrical rays, A ray is the path of propagation of energy in the zero wavelength limit ($\lambda \rightarrow 0$). The way in which rays may represent the propagation of wavefronts lor some familiar situations is shown in Fig. 2.5. You will note that a plane wavefront corresponds to parallel rays and spherical wavefronts correspond to rays diverging from a point or converging to a point. You will agree that all parts of the wavefront take the same time to travel from the source.

Huygens proposed that light propagates as a wavefront (a surface of constant phase) progresses in a medium perpendicular to itself with the speed of light. The zero wavelength approximation of wave optics is known as **geometrical optics.**

The laws of geometrical optics are incorporated in Fermat's principle. We will now discuss it in detail.

2.4 FERMAT'S PRINCIPLE

In its original form, Fermat's principle may be stated as follows;

> **Any light ray travels between two end points along a line requiring the minimum transit time.**

If v is the speed of light at a given point in a medium, the time taken to cover the distance dl is

$$dt = \frac{dl}{v} \tag{2.22}$$

In your earlier years you have learnt that the refractive index of a medium is defined as the ratio of the speed of light in vacuum to its speed in the medium, i.e.

$$n = \frac{c}{v}$$

Using this relation in Eq. (2.22), we get

Hence, the time taken by light in covering the distance from point A to B is

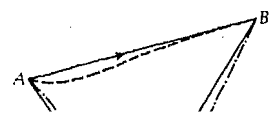
$$\tau = \frac{1}{c}\int_A^B n\, dl$$

The quantity

$$L = \int_A^B n\, dl \tag{2.23}$$

has the dimensions of length and is called the **optical distance** or **optical path length** between two given points. You must realise that optical distance is different from the physical (geometrical) distance ($= \int_A^B dl$). However, in a homogeneous medium, the optical distance is equal to the product of the geometrical length and the refractive index of the medium. Thus, we can write

$$\tau = \frac{L}{c}$$



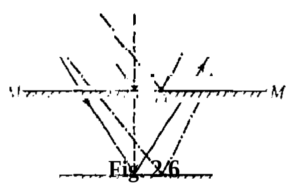


Fig. 2.6

This is Fermat's principle of least time. Let us pause for a moment and ask: Is there any exception to this law? Yes, there are cases where the optical path corresponds to maximum time or it is neither a maximum nor a minimum, i.e. stationary. To incorporate such situations, this principle is modified as follows:

Out of many paths connecting two given points, the light ray follows that path for which the time required is an extremum. In other words, the optical path length between any two points is a maximum, minimum, or stationary.

The essential point involved in Fermat's principle is that slight variation in the actual path causes a second-order variation in the actual path. Let us consider that light propagates from point A in the medium characterised by the refractive index n to the point B as shown in Fig. 2.0. According to this principle,

$$\delta \int_A^B n(x, y, z)\, dl = 0 \tag{2.24}$$

For a homogeneous medium, the rays are straight lines, since the shortest optical path between two points is along a straight line.

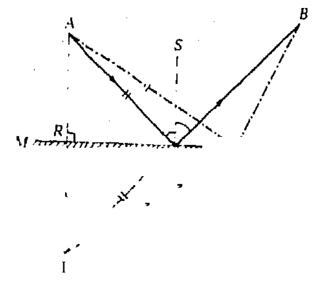


Fig. 2.7 : Reflection of rays at plane interface

In effect, Fermat's principle prohibits the consideration of an isolated ray of light. It tells us that a path is real only when we extend our examination to the paths in **immediate neighbourhood** of the rays. To understand the meaning of this statement, let us consider the case of finding the path of a ray from a point A to a point B when both of them lie on the same side of a mirror M (Fig. 2.7). It can be seen that the ray can go directly from A to B without suffering any reflection. Alternatively, it can go along the path APB after suffering a single reflection from the mirror. If Fermat's principle had asked for, say, an absolute minimum, then the path APB would be prohibited; but that is not the actual case. The path APB is also minimum in the neighbourhood involving paths like AQB. The phrase "immediate neighbourhood of path" would mean those paths that lie near the path under consideration and are similar to it. For example, the path AQB lies near APB and is similar to it; along both paths the ray suffers one reflection at the mirror. Thus Fermat's principle requires an extremum in the immediate neighbourhood of the actual path, and in general, there may be more than one ray path connecting two points.

All the laws of geometrical optics are incorporated in Fermat's principle. We now illustrate Fermat's principle by applying it to reflection of light.

Example 1

Using Fermat's principle, derive the laws of reflection.

Solution

Let us first consider the case of reflection. Refer to Fig. 2.8. Light from a point A is reflected at a mirror MM towards a point B. A ray APB connects A and B. θ_I and θ_R are the angles of incidence and reflection, respectively. We have denoted the vertical distances of A and B from the mirror MM by a and b. From the construction in Fig. 2.8 and Pythagoras' theorem, we find that the total path length I of this ray from A to MM to B is

$$l = \sqrt{a^2 + x^2} + \sqrt{b^2 + (d - x)^2} \qquad (2.25)$$

wherd x is the distance between the foot of the perpendicular from A and the point P at which the ray touches the mirror.

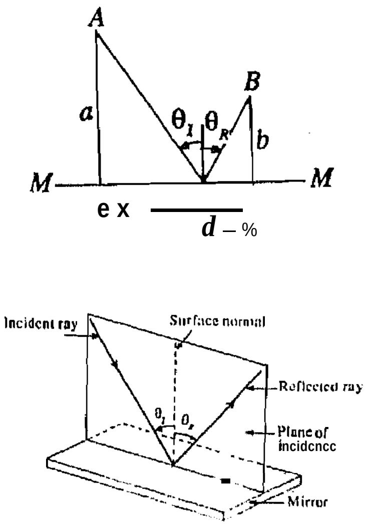


Fig. 2.8 Derivation of the laws of reflection using Fermat's principle.

According to Fermat's principle, P will have a position such that the time of travel of the light must be a minimum (a maximum or stationary), Expressed in another way, the total length I of the ray must be a minimum or maximum or stationary, In other words, for Fermat's principle to hold, the derivativc of l with respect to x must be zero, i.e. $dl/dx = 0$. Hence, on differentiating Eq. (2.25) with respect ton, we get

$$\frac{}{dx} \quad \frac{1}{2}(a^2 + x^2)^{-1/2}(2x) + \frac{1}{2}\left[b^2 + (d - x)^2\right]^{-1/2} \times 2(d - x)(-1) = 0 \qquad (2.26)$$

which can be rewritten as

$$\frac{x}{(a^2 + x^2)^{1/2}} = \frac{d - x}{[b^2 + (d - x)^2]^{1/2}} \qquad (2.27)$$

By examining Fig. (2.8) you will not that this gives

$$\sin\theta_I = \sin\theta_R$$

or

$$\theta_I = \theta_R \qquad (2.28)$$

which is (part of) the law of reflection. You will also note that the incident ray, the reflected ray and the normal to MM lie in the same incidence plane.

In the above example time required or the optical palh length can be seen to be minimum hy calculating the second deviative and finding its value at x for which $dl/dx = 0$. The 2nd derivative turns out to bc positive, showing it to be minimum. You can convince, yourself by carrying out this simple caculation.

We now summarise what you have learnt in this unit.

2.5 SUMMARY

- When an e.m. wave is incident normally an the interface separating two optically different media. the reflected and transmitted electric field amplitudes are given by

$$E_{0R} = \frac{1-\alpha}{1+\alpha} E_{0I}$$

and

$$E_{0T} = \frac{2}{1+\alpha} E_{0I}$$

where $\alpha = \sqrt{\mu_1 \varepsilon_1 / \mu_2 \varepsilon_2}$ and E_{0I} is amplitude of incident electric field.

- The frequency of an e.m. wave is not affected when it undergoes reflection or refraction.
- Fermat's principle states that a ray of light travels between two given points along that path for which the time required is an extremum:

$$\delta \int_A^B n\,(x, y, z)\, dl = 0$$

2.6 TERMINAL QUESTIONS

1. Derive Snell's law from Fermat's principle.
2. A collimated beam is incident parallel to the axis of a concave mirror. It is reflected into a converging beam. Using Fermat's principle show that the mirror is parabolic.

2.7 SOLUTIONS AND ANSWERS

TQs

1. To prove the law of refraction from Fermat's principle, consider Fig. 2.9, which shows that the points *A* and B are in two optically different media. (If the refractive index on both sides of the boundary *SS* were the same, the path from *A* to *B* would be a straight line, irrespective of the magnitude of the refractive index. But the refractive indices are not the same and the ray APB is not a straight line.) Suppose that the velocities of light on the two sides of the boundary are v_1 and v_2. Since $v = l/t$, the time light takes to traverse the paths AP and PB is

$$t = \frac{\sqrt{a^2 + x^2}}{v_1} + \frac{\sqrt{b^2 + (d-x)^2}}{v_2} = \frac{l_1}{v_1} + \frac{l_2}{v_2} \qquad \text{(i)}$$

Using the relation $n = c/v$, this can be rewritten as

$$t = \frac{n_1 l_1 + n_2 l_2}{c} = \frac{l}{c}$$

where $l\ (= n_1 l_1 + n_2 l_2)$ is the optical path length of the ray. The geometrical path in this case is $l_1 + l_2$. If λ is the wavelength of light in vacuum and λ_m in a medium of refractive index n, then $\lambda = n\lambda_m$. This shows that the optical path length is equal to the length that the same number of waves would have if the medium were a vacuum.

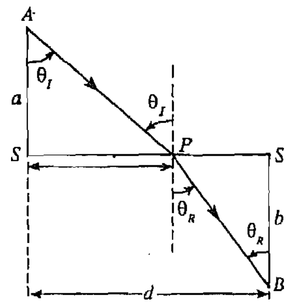


Fig.2.9: A ray from A passes to B after refraction at P

Fermat's principle requires that $dl / dx = 0$ for some values of x. The optical path length

$$l = l_1 n_1 + l_2 n_2$$

$$= n_1 \sqrt{(a^2 + x^2)} + n_2 \sqrt{b^2 + (d - x)^2} \qquad \text{(ii)}$$

so that $$\frac{dl}{dx} = n_1 \frac{x}{\left(a^2 + x^2\right)^{1/2}} - n_2 \frac{d - x}{\left\{b^2 + (d - x)^2\right\}^{1/2}} = 0$$

or $$n_2 \frac{x}{\left(a^2 + x^2\right)^{1/2}} = n_2 \frac{d - x}{\left\{b^2 + (d - x)^2\right\}^{1/2}} \qquad \textbf{(iii)}$$

As before, we can write it in terms of the angles of incidence and refraction as

$$n_I \sin \theta_I = n_2 \sin \theta_R \qquad \text{(iv)}$$

which is **Snell's law of refraction.** It shows that when light passes from a medium of lower refractive index (rarer medium) to a medium of higher refractive index (denser medium), it bends towards the surface normal.

2

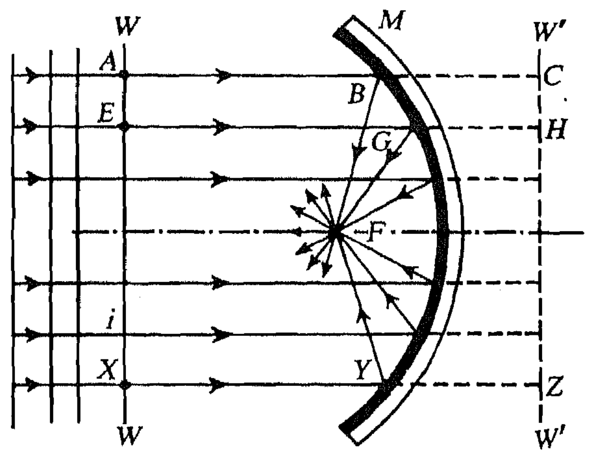


Fig. 2.10: Reflection of light incident on a concave mirror

Fig. 2.10: depicts cross sectional view of parallel rays corresponding to a plane wave WW incident on the mirror. The reflected rays converge on F. The optical path lengths of all rays reaching F must be the same:

$$n_i (\overline{AB} + \overline{BF}) = n_i (\overline{EG} + \overline{GF}) = \ldots n_i (\overline{XY} + \overline{YF})$$

Now let the line segments $\overline{AB}$, $\overline{EG}$, ... , $\overline{XY}$ be prolonged through the mirror to points C, H, ,.., Z such that

$$\overline{BC} = \overline{BF}, \quad \overline{GH} = \overline{GF}, \ldots, \overline{YZ} = \overline{YF}$$

The two sets of equalities above imply that $\overline{AB} + \overline{BC} = \overline{EG} + \overline{GH} = \ldots$ $= \overline{XY} + \overline{YZ}$, which tells us that the distance between WW and $W'W'$ through C, H, ..., Z is constant. We have thus constructed a straight line W' W' such that the points of M are equidistant from it and point F. By definition, then M is parabolic (with *Focus F*).