Uttar Pradesh Rajarshi Tandon Open University

Bachelor of Science

DCEPHS-105

OPTICS

DCEPHS-105

OPTICS

Uttar Pradesh Rajarshi Tandon Open University

Bachelor of Science

Block

1 Geometrical and Quantum Optics

UNIT- 1 CO-AXIAL SYSTEM OF LENSES

1.17 Suggested Readings

1.1 INTRODUCTION:

- The branch of physics called optics deals with the behaviour of light and other electromagnetic waves. Light is the principal means by which we gain knowledge of the world. Consequently, the nature of light has been the source of one of the longest debates in the history of science.
- Electromagnetic radiation with wavelengths in the range of about 4000 A to 7000 A, to which eye is sensitive is called light.
- In order to understand the nature of light, various theories have been given from time to time. The first scientific theory, to explain the nature of light was proposed by Newton in 1675. According to this theory, light is a stream of tiny particles of negligible mass. These tiny particles, called corpuscles, travel through a medium at high speed and cause the sensation of vision upon entering the eye. This theory could explain reflection and rectilinear propagation of light, but fails to explain the phenomena like interference, diffraction and polarization of light. Therefore, this theory was rejected. Second theory that explains the nature of light is Huygen's wave theory. According to this theory, light is transmitted from one point to another in the form of waves. This theory could successfully explain the basic laws of reflection and refraction of light. It could also explain the phenomenon like interference, diffraction, etc. Huygen's wave theory proposed the existence of luminiferous ether medium. Light travels ether in the form of waves. But experimentally, it was found that there is no existence of such as medium.
- Later on, James Clerk Maxwell put forth electromagnetic theory, according to which light waves are electromagnetic waves which do not require any material medium for its

propagation. The main drawback of Huygen's wave theory is that it could not explain the phenomenon like photoelectric effect, Compton effect, etc.

- In 1905, Einstein interpreted the photoelectric effect by putting forward his photon theory, according to which the energy associated with light beam of frequency v is concentrated in corpuscles of energy hv, where h is Planck's constant. The experiment like photoelectric effect, proves that light has particle properties. On the other hand, the phenomena like interference, diffraction demonstrate that light has wave-like properties. But with De-Broglie's hypothesis of matter waves, the problem of dual behaviour of light was resolved. From the point of view of quantum mechanics, light is just like a wave or just like a particle depending upon the experiment. So, we think of light as both a wave and a particle. The wave-particle nature of light is evident in the formation of optical images.

  The subject of optics is studied under the following three heads:

  (a) Geometrical optics,

  (b) Physical optics and

  (c) Quantum optics

- The first group, geometrical optics, deals with geometrical formation of images by mirrors, lenses and prism. In the study of geometrical optics it is assumed that light travels through homogeneous, isotropic media in straight lines and hence path of light rays is represented by well-defined geometric lines. It uses the concept of rays. This assumption is not strictly valid because light is observed to bend round the corners of an obstacle, whose dimensions are comparable with the wavelength of light. Thus, the assumption that light travels in a straight line in homogeneous isotropic media should, therefore be considered with limitations.

- The fundamental laws which form the basis of geometrical optics are :
  - (a) rectilinear propagation of light,
  - (b) laws of reflection and refraction of light.
- The present chapter is entirely based on geometrical optics. So, in this chapter, light is represented by rays which are geometrical lines along which light flows.
- The second group, physical optics covers the phenomena of interference, diffraction, polarization and double refraction, etc. In physical optics, above phenomena ae described on the basis of wave nature of light.
- The third group, quantum optics treats light as a particle called photons. This group covers the phenomena like photoelectric effect, Compton effect.

1.2 OBJECTIVES

After studying this Unit, students will be able to:

- Explain the concept of convex lens and concave lens.
- Define radius of Curvature, Poles of Lens.
- Discuss the Sing Convention of the lens.
- Understand the Concept of Eye Pieces
- Difference between Huygen’s and Ramsdon’s Eye Pieces.
- Explain the concept of Cardinal Points.
- Understand the concept of Aplanatic Points.

1.3 GEOMETRICAL AND QUANTUM OPTICS

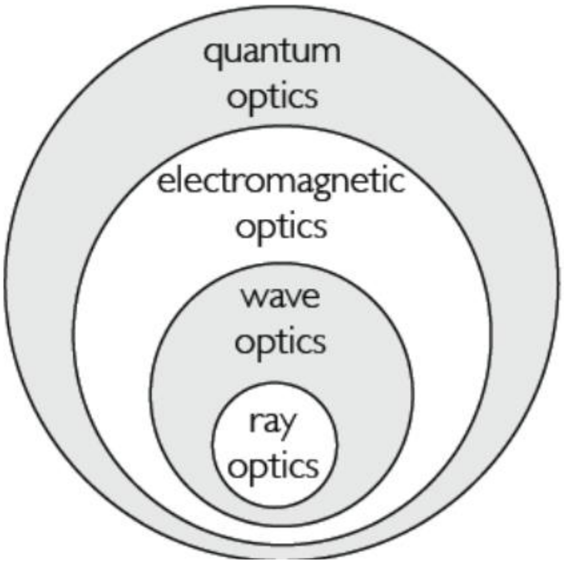
Light is a form of radiant energy, that is, energy emitted by excited atoms or molecules which can cause the sensation of vision in a normal human eye.

The branch of physics which deals with the phenomena of light is called optics. The three branches of Optics are as follows:

Geometrical Optics: This deals with the study of light in which light is considered as moving along a straight line as a ray. A ray of light gives the direction of propagation of light. When light meets a surface which separates two media, reflection and refraction take place. An image or an array of images may be formed due to this.

Physical Optics: It deals with the theories regarding the nature of light and provides an explanation for the different phenomena in light, such as reflection, refraction, interference, diffraction, polarization, and rectilinear propagation.

Quantum optics: It is the study of how individual quanta of light, known as photons, interact with atoms and molecules. This includes studying the particle-like properties of photons. Photons have been used to test many of the counter-intuitive predictions of quantum mechanics, such as entanglement and teleportation, and are a useful resource for quantum information processing.



1.4 LENSES: THIN AND THICK

Lenses: A lens is a portion of transparent material bounded by two curved surfaces or by one curved and other plane surface. The curved surfaces of commonly used lenses are mostly spherical, however, sometimes lenses having cylindrical surfaces are used for specific purposes. In the present unit, we shall study only the lenses having spherical surfaces.

Lenses can be classified into two groups:

(a) Converging or convex lenses and

(b) Diverging or concave lenses.

1.4.1 Converging or Convex Lenses

If a parallel beam of light is incident on a convex lens, all the rays, after refraction through the lens, approach one another and converge at a point as shown in Fig. 1

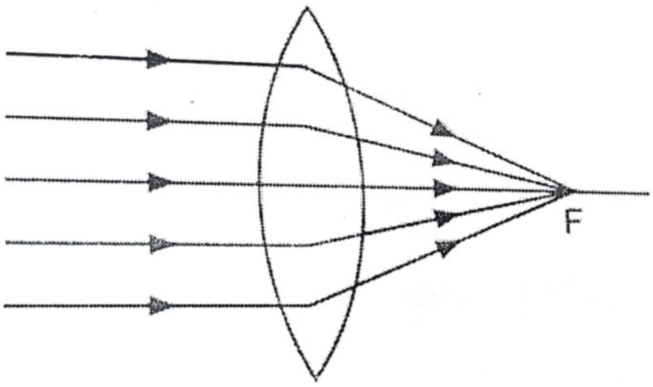


Figure: 1

There are three types of convex lenses.

(i) **Bi-convex or Double convex lens:** In this type, both surface of a lens are convex.

(ii) **Plano-convex lens:** In this type of lens, one surface is convex and the other surface is plane.

(iii) **Concavo-convex lens:** In this type, one surface is convex, while the other surface is concave.

Fig. 2 shows the three types of convex lenses.

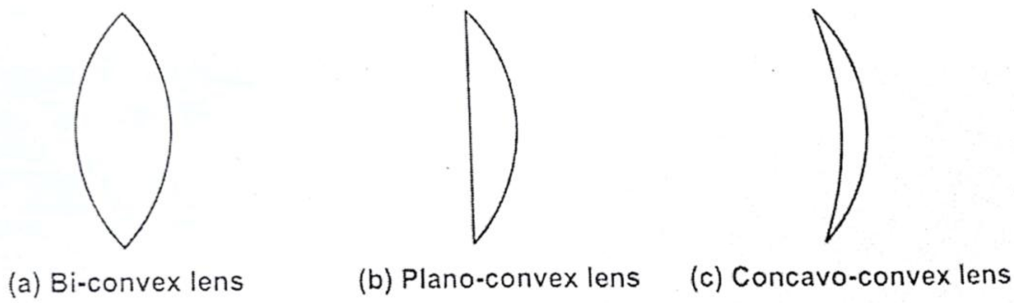


Figure: 2

The distinguishable feature of any type of convex lens is that it is thicker at its centre than at its rim or periphery.

1.4.2 Diverging or Concave Lenses

When a parallel beam of light is incident on such a lens, all the rays after refraction move away from one another as shown in Fig. 1.2. In the figure, all the rays appear to be diverging from a single point.

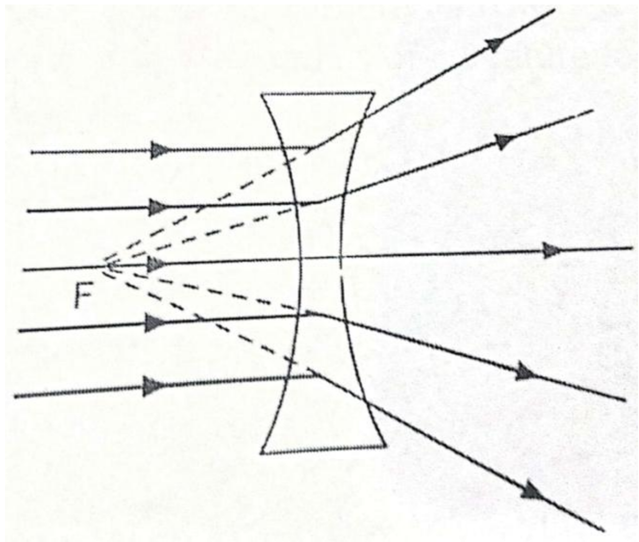


Figure: 3

There are three types of concave lenses.

(i) **Bi-concave or Double convex lens:** In this type, both surface of a lens is concave.

(ii) **Plano-concave lens:** In this type, one surface is concave and the other surface is plane.

(iii) **Convexo-concave lens:** In this type, one surface of a lens is convex, while the other surface is concave.

Fig. 4 shows three types of concave lenses.

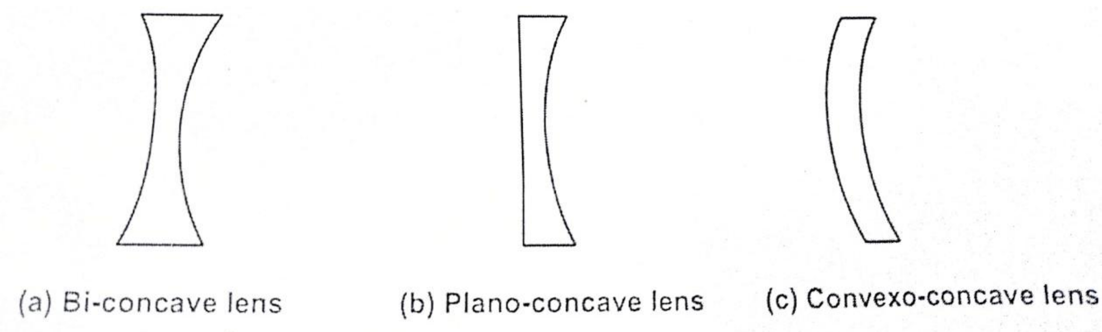


Figure: 4

The distinguishable feature of any type of concave lens is that it is thinner at its centre than at its rim periphery.

1.4.3 Centre of Curvature (C_1 and C_2)

There are two centres of curvature for a lens, one each belonging to both the surfaces.

The centre of curvature of a surface of a lens is the centre of that sphere of which the surface forms a part. Centres of curvature for a concave and convex lens are shown in Fig. 1.5 (a) and (b) respectively.

1.4.4 Radius of Curvature (R_1 and R_2)

There are two radii of curvature of a lens, one each belonging to both the surface. The radius of curvature for concave and convex lens is shown in Fig. 5 (a) and (b) respectively.

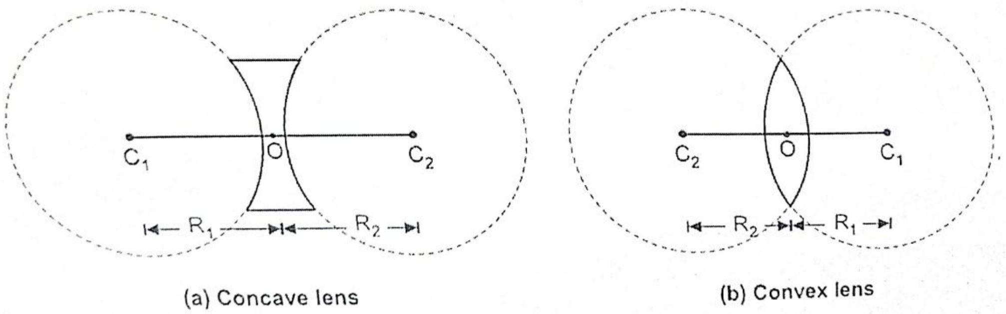


Figure: 5

1.4.5 Axis of the Lens

A line passing through the centres of curvature of both the faces of a lens is called the axis of the lens. A section of the lens perpendicular to the axis of the lens is called principal plane of the lens.

1.4.6 Optical Centre (C)

The point of intersection of principal plane with the axis of the lens is called optical centre of the lens.

1.4.7 Poles of the Lens

The points at which the axis of the lens cut the surfaces of the lens is called poles of the lens.

1.4.8 Principal Focus (F)

Consider a parallel beam of light coming parallel to the principal axis of the lens. In case of convex lens, the beam converges and meets at a point F on the principal axis after refraction through the lens. In case of concave lens, the beam diverges and appears to come from point F after refraction through the lens. The point F in Fig. 6 (a) and (b) is called the principal focus of the lens.

In short, principal focus of a lens is a point at which a beam of light coming parallel to principal axis meets or appears to meet after refraction through the lens.

The distance of the principal focus from the optical centre of the lens is called the focal length (f) of the lens.

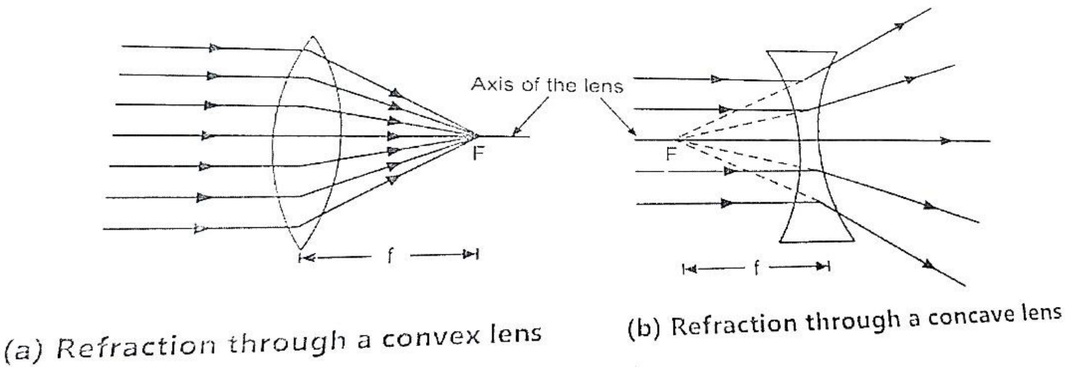


Figure: 6 (a) & (b)

1.5 SIGN CONVENTIONS

To study the formation of image using a lens, it is essential to specify distances, heights of objects and images from certain reference points. Different sign conventions are used in practice. We shall follow the easiest convention of signs i.e., co-ordinate geometry because we are already familiar with this convention. The widely followed new cartesian sign conventions are given below.

(i) All figures are drawn such that the incident ray travels from left to right and object should be placed to the left of the lens.

(ii) All distances are measured from the optical centre of the lens. The pole or optical centre is chosen as origin. In Fig. 7, O is the pole or optical centre of the lens.

(iii) Distances measured to the left of the pole or optical centre are negative and those to the right are positive.

(iv) Distances measured above the axis are positive, while those measured below the axis are negative. Above conventions of signs are represented in Fig. 7 .

(v) The angle which a ray makes with the axis of the lens is taken as positive when the axis is to the rotated through an acute angle in anticlockwise direction to make it coincide with the ray.

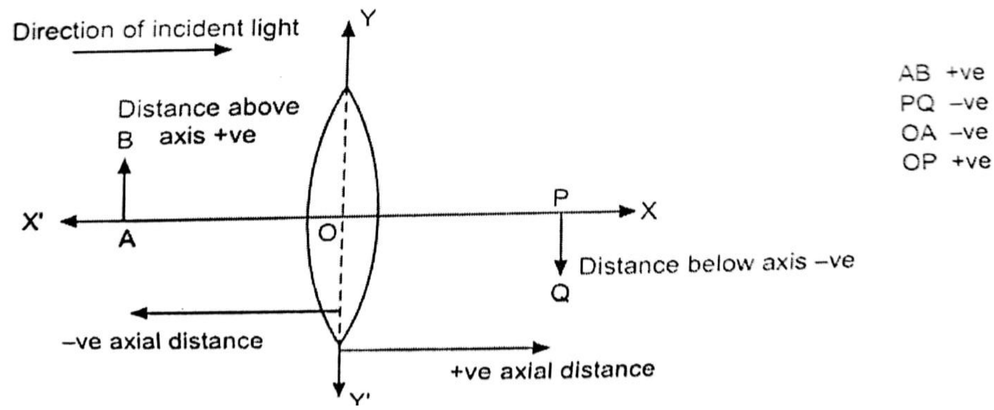


Figure: 7

According to this sign convention, R_1 is negative and R_2 is positive for concave lens [Fig. 5 (a)]; while R_1 is positive and R_2 is negative for convex lens [Fig. 5 (b)]. The focal length of a convex lens is positive [Fig. 6 (a)]; while the focal length of a concave lens is negative [Fig. 6 (b)].

1.5.1 Thin Lenses: Lens Equation

It is the equation connecting focal length of the lens with distances of objects and images.

To obtain lens equation consider a thin convex lens of focal length f as shown in Fig. 8. Let an object AB be situated on the principal axis of a lens. Consider two rays AC and AD. The ray AD goes parallel to principal axis and after refraction it passes through focus F. The ray AC passes straight through optical centre (C). Since the object AB is situated beyond F of a convex lens, a real image A_1B_1 is formed as shown in Fig. 8.

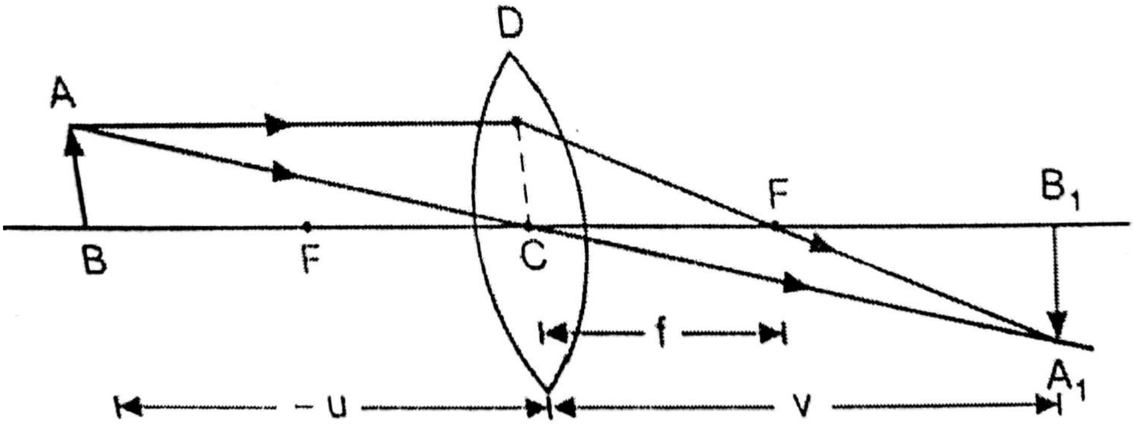


Figure: 8

In Fig. 8, Δ ABC and Δ A_1B_1C are similar triangles.

∴ $$\frac{AB}{A_1B_1} = \frac{BC}{B_1C} \qquad \text{............ (1)}$$

Δ CDF and Δ A_1B_1F are also similar triangles.

∴ $$\frac{CD}{A_1B_1} = \frac{CF}{B_1F}$$

But, CD = AB

$\therefore \quad \frac{AB}{A_1B_1} = \frac{CF}{B_1F}$ ………. (2)

From equations (1) and (2), we get

$$\frac{BC}{B_1C} = \frac{CG}{B_1F}$$

But in Fig, s8, $B_1f = B_1C - CF$

$\therefore \quad \frac{BC}{B_1C} = \frac{CF}{B_1C - CF}$

According to sign conventions, we have

$BC = -u$, $B_1C = v$ and $CF = f$

$\therefore \quad \frac{-u}{v} = \frac{f}{v-f}$

or $-u(v - f) = f v$

$-uv + uf = f v$

Dividing throughout above equation by uvf, we get

$\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$ …………. (3)

This equation is called lens formula.

1.5.2 Lens Maker's formula for a thin lens

If a thin lens made up of material of refractive index μ_2 is kept in the medium of refractive index μ_1, then its focal length in the given medium is given by

$$\frac{1}{f} = \left(\frac{\mu_2}{\mu_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

Or $\frac{1}{v} - \frac{1}{u} = \left(\frac{\mu_2}{\mu_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$

Where, R_1 is the radius of curvature of first surface of the lens, R_2 is the radius of curvature of second surface of the lens and f is focal length of the lens. If the lens made up f material of R.I. μ is kept in air, then the focal length of the lens in air medium is given by

$$\boxed{\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)} \qquad \ldots\ldots\ldots\ldots\ldots (4)$$

1.5.3 Equivalent focal length

Equivalent focal length of two thin lenses placed co-axially and separated by a distance x from each other is given by

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{x}{f_1 f_2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots.. (5)$$

or $$\boxed{f = \frac{f_1 f_2}{f_1 + f_2 - x} = \frac{f_1 f_2}{\Delta}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

where, $\Delta = f_1 + f_2 - x$ is called the optical separation or optical interval between two thin lenses, f_1 is focal length of first lens and f_2 is focal length of second lens.

The equivalent focal length of two thin lenses in contact is given by

$$\boxed{\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

1.6 WHAT IS A MIRROR?

A mirror is a reflective surface that bounces off light, producing either a real image or a virtual image. When an object is placed in front of a mirror, the image of the same object is seen in the mirror. The object is the source of the incident rays and the image is formed by the reflected rays. Based on the intersection of light rays, the images are classified as either a real images or a virtual image. A real image occurs when the light rays actually intersect while virtual images occur due to the apparent divergence of light rays from a point.

Ray diagrams help us trace the path of the light for the person to view a point on the image of an object. The Ray diagram uses lines with arrows to represent the incident and reflected ray. It also helps us trace the direction in which the light travels.

Concave Mirrors and Convex Mirrors:

The surface that reflects nearly every type of light that is incident on it is known as a mirror. A mirror can have a flat surface or a curved surface. A mirror with a flat surface is called a plane mirror, and a mirror with a curved surface is called a spherical mirror. In this article, let us learn about convex mirrors and concave mirrors.

1.6.1 Plane Mirror vs Spherical Mirrors

Mirrors are made into different shapes for different purposes.

The two of the most prominent types of mirrors are:

- Plane Mirrors
- Spherical Mirrors

A plane mirror is a flat, smooth reflective surface. A plane mirror always forms a virtual image that is upright, and of the same shape and size as the object, it is reflecting. A spherical mirror is a mirror that has a consistent curve and a constant radius of curvature. The images formed by a spherical mirror can either be real or virtual. **Spherical mirrors are of two types as:**

  - Concave Mirror
  - Convex Mirror

In the next few sections, let us learn in-depth about the characteristics of convex and concave mirrors and the images formed by them when the object is kept at different positions.

Spherical Mirrors

Spherical mirrors are mirrors having curved surfaces that are painted on one of the sides. Spherical mirrors in which inward surfaces are painted are known as convex mirrors, while the spherical mirrors in which outward surfaces are painted are considered concave mirrors.

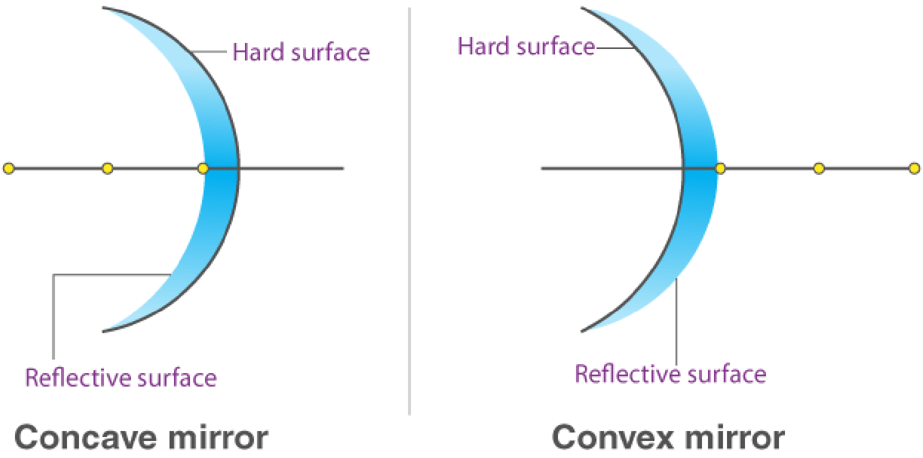


Figure: 9

1.6.2 Concave Mirror

If a hollow sphere is cut into parts and the outer surface of the cut part is painted, then it becomes a mirror with its inner surface as the reflecting surface. This type of mirror is known as a concave mirror.

1.6.3 Characteristics of Concave Mirrors

- Light converges at a point when it strikes and reflects back from the reflecting surface of the concave mirror. Hence, it is also known as a converging mirror.
- When the concave mirror is placed very close to the object, a magnified, erect and virtual image is obtained.
- However, if we increase the distance between the object and the mirror then the size of the image reduces and a real and inverted image is formed.

- The image formed by the concave mirror can be small or large and can be real or virtual.

1.6.4 Convex Mirror

If the cut part of the hollow sphere is painted from the inside, then its outer surface becomes the reflecting surface. This kind of mirror is known as a convex mirror.

1.6.5 Characteristics of Convex Mirrors

- A convex mirror is also known as a diverging mirror as this mirror diverges light rays when they strike its reflecting surface.
- Virtual, erect, and diminished images are always formed with convex mirrors, irrespective of the distance between the object and the mirror.

1.6.6 Image Formation by Spherical Mirrors

Guidelines for Rays Falling on the Concave and Convex Mirrors

1. When a ray strike concave or convex mirrors obliquely at its pole, it is reflected obliquely making the same angle with the principal axis.
2. When a ray, parallel to the principal axis strikes concave or convex mirrors, the reflected ray passes through the focus on the principal axis.
3. When a ray, passing through focus strikes concave or convex mirrors, the reflected ray will be parallel to the principal axis.
4. A ray passing through the centre of curvature of the spherical mirror will retrace its path after reflection.

1.6.7 Image Formation by Concave Mirror

By changing the position of the object from the concave mirror, different types of images can be formed. Different types of images are formed when the object is placed:

1. At the infinity
2. Beyond the centre of curvature
3. At the centre of curvature
4. Between the centre of curvature and principal focus
5. At the principal focus
6. Between the principal focus and pole

1.6.8 Concave Mirror Ray Diagram

- Concave Mirror Ray Diagram lets us understand that, when an object is placed at infinity, a real and inverted image is formed at the focus. The size of the image is much smaller compared to that of the object.

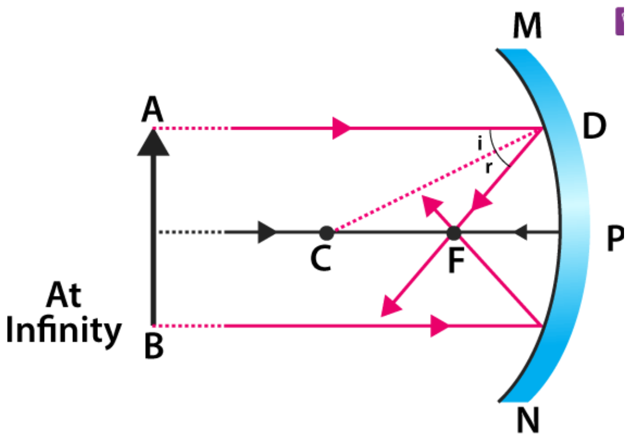


Figure: 10

When an object is placed behind the center of curvature, a real image is formed between the center of curvature and focus. The size of the image is smaller than compared to that of the object.

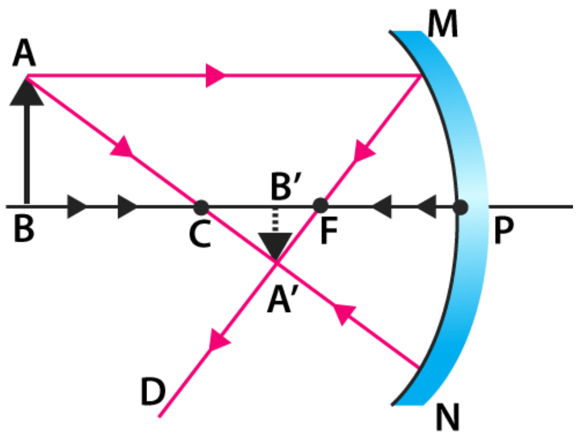


Figure: 11

- When an object is placed at the center of curvature and focus, the real image is formed at the center of curvature. The size of the image is the same as compared to that of the object.

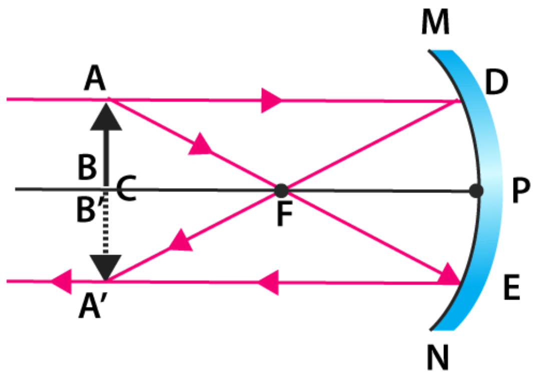


Figure: 12

- When an object is placed in between the center of curvature and focus, the real image is formed behind the center of curvature. The size of the image is larger than compared to that of the object.

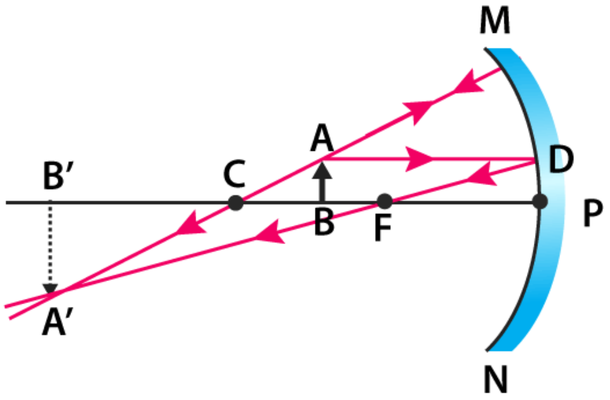


Figure: 13

- When an object is placed at the focus, the real image is formed at infinity. The size of the image is much larger than compared to that of the object.

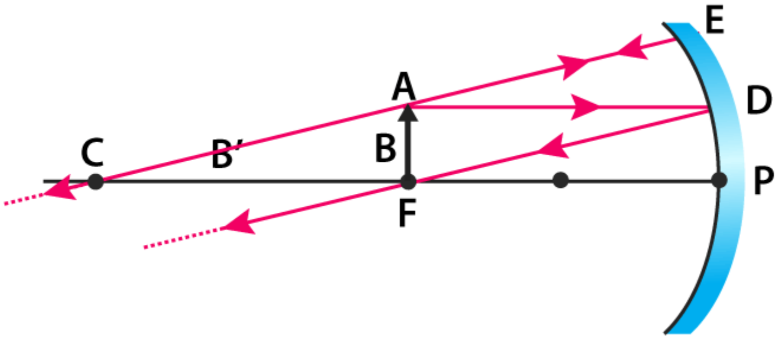


Figure : 14

- When an object is placed in between focus and pole, a virtual and erect image is formed. The size of the image is larger than compared to that of the object.

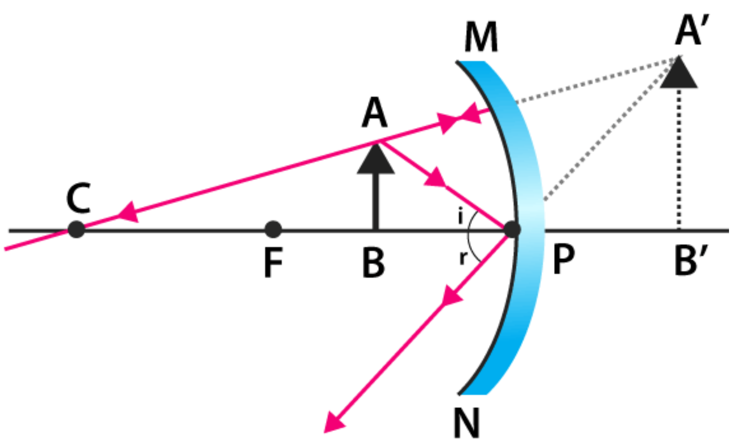


Figure: 15

1.6.9 Image Formation by Convex Mirror

The image formed in a convex mirror is always virtual and erect, whatever be the position of the object. In this section, let us look at the types of images formed by a convex mirror.

- When an object is placed at infinity, a virtual image is formed at the focus. The size of the image is much smaller than compared to that of the object.

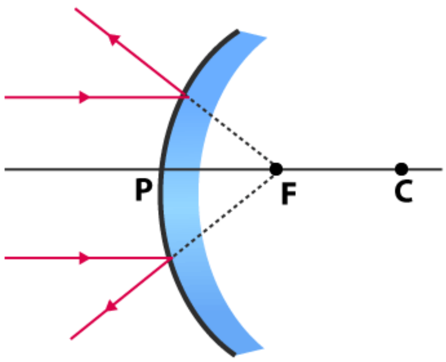


Figure: 16

- When an object is placed at a finite distance from the mirror, a virtual image is formed between the pole and the focus of the convex mirror. The size of the image is smaller than compared to that of the object.

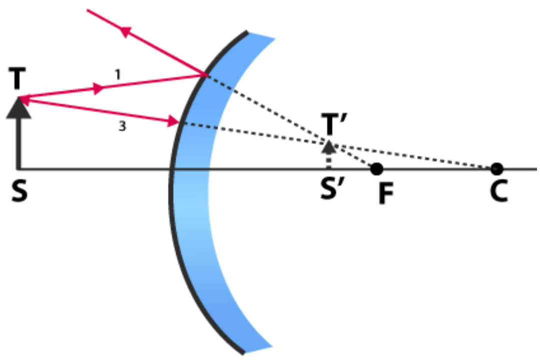


Figure: 17

1.7 EYE – PIECES

- When a single convex lens is used as an eye - piece, the rays coming from the peripheral parts of the image of an object formed by the object would cross over the eye lens and fail to the enter the eye which is kept close to the eye lens. This reduces the field of view.
- To overcome this defect, another lens called field lens is introduced between the eye lens and the objective.

1.7.1 Huygen's & Ramsden Eye-Pieces

Huygen's eye-piece		**Ramsden's eye-piece**	
1.	The image formed by the objective lies between field lens and eye lens.	1.	The image formed by the objective lies in front of the field lens.
2.	No cross-wires can be used.	2.	Cross-wires can be used.
3.	The condition for minimum spherical aberration and achromatism is satisfied.	3.	The conditions for minimum spherical aberration and achromatism are not satisfied, but can be made achromatic by using achromatic doublet.
4.	The image formed is slightly convex towards the eye and hence the other aberrations are not completely eliminated	4.	The image formed is flat and hence the other aberrations are completely eliminated.
5.	Used in microscopes and other optical instruments for qualitative observations only.	5.	Used in telescopes and other optical instruments where accurate measurements are required.
6.	The distance between the eye lens	6.	The distance between eye lens and the

	and the eye is too small and hence more strain to the eye.		eye and the eye is greater and hence less strain to the eye.

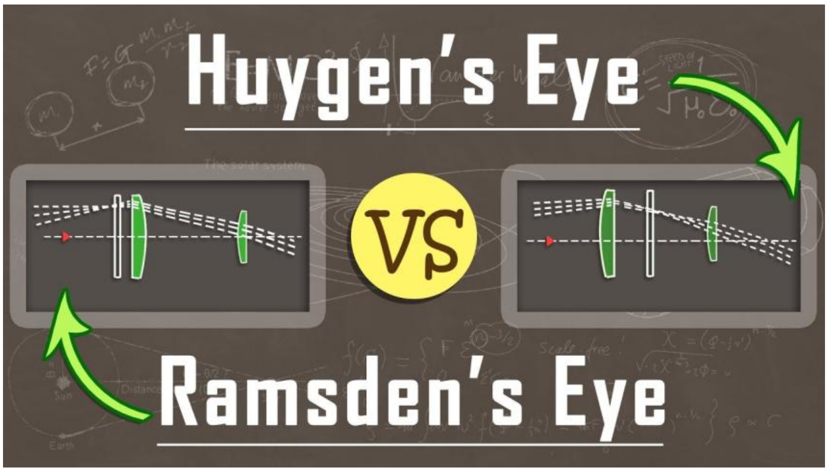


Figure: 18

1.8 CARDINAL POINTS

- In the optical instruments a single lens is rarely used, but two or more lenses separated from each other are generally used. In order to determine the position and size of the image formed by such a complex optical system, one must have knowledge of cardinal points. A combination of lenses having common principal axis is called the co-axial system of lenses.
- We have seen that, in case of refraction through a single thin lens, the thickness of the lens is neglected while calculating various distances. When a light gets refracted through a thin lens, it is immaterial from which point of the lens, we measure the distances. Whereas in case of a thick lens or a co-axial system of lenses, we cannot proceed with this assumption. It is tedious to determine the position and size of the image formed by thick lens or a co-axial system of lenses by considering the

refraction at each surface. In order to overcome this difficulty, Gauss in 1841 proved that any number of co-axial refracting systems can be replaced by a single thin lens.

- Thin lens formula can be then applied with the restriction that the object and image distance are measured from two theoretical (imaginary) planes fixed with reference to the refracting system. Formation of images due to thick lens can also be studied by this method.
- The points of intersection of these planes with the axis of an optical system are called the principal points or Gauss points. Actually, there are six points in all, called as cardinal points of an optical system with the help of which, the position and size of the image of a given object can be found. The following are the cardinal points in an optical system.

(a) Two Focal Points (F_1 and F_2)

(b) Two Principal Points (P_1 and P_2)

(c) Two Modal Points (N_1 and N_2)

1.8.1 Focal Points and Focal Planes

- Consider an optical system consisting of a thick convex lens as shown in Fig.19.
- Consider as set of rays parallel to the principal axis incident on the lens (shown in Fig. by →-). After refraction, these rays converge to a point F_2 on the principal axis. The point F_2 is called second focal point or second principal focus and is the position of the image corresponding to the axial point object situated at infinity (i.e., $u = \infty$).
- Similarly, the rays starting from an axial point F_1, after refraction through the lens (shown in Fig. 19 by -→) become parallel to the principal axis. The point F_1 is called first principal focus or first focal point and is the position of object whose

image is formed at infinity (i.e. $v = \infty$). The points F_1 and F_2 are called principal foci or focal points of the convex lens.

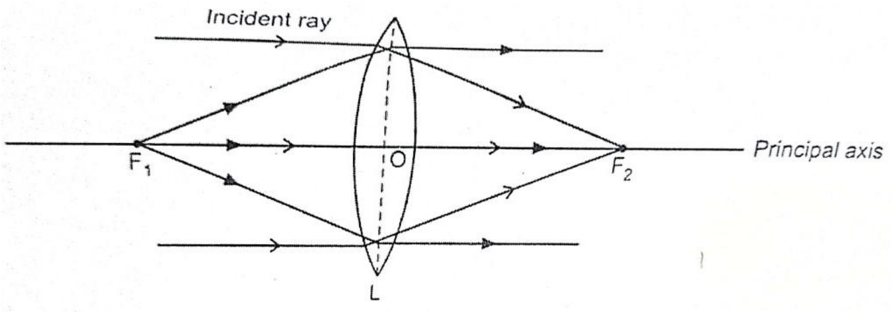


Figure: 19

- Fig. 20 shows two focal points F_1 and F_2 for a thick concave lens. Consider a set of rays parallel to the principal axis incident on the concave lens (shown in Fig. 20 by -→). After refraction through the lens, the emergent rays appear to diverge from a point F_2 on the principal axis. The point F_2 is called second focal point of the concave lens.
- Similarly, the rays directed towards an axial point F_1 (shown in Fig. 20 by (-→), after refraction through the lens becomes parallel to the principal axis. The point F_1 is called first focal point of the concave lens.
- The two points F_1 and F_2 are called principal foci or focal points of the concave lens.

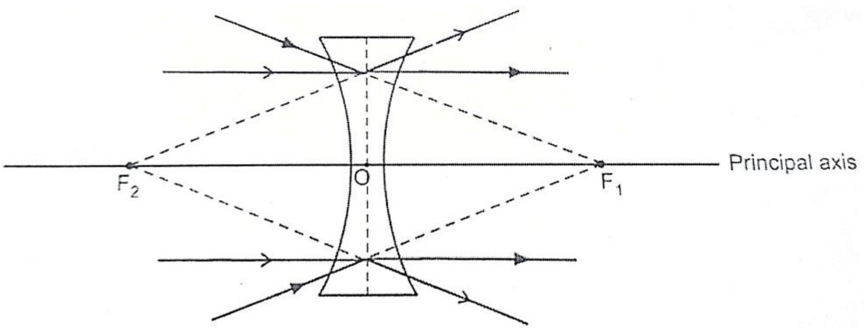


Figure: 20

- The planes passing through the focal points and perpendicular to the principal axis are called focal planes. The important property of the focal planes is that the rays starting from any point on the focal plane in the object space correspond to a set of conjugate parallel rays in the image space. Similarly, a set of parallel rays in the object space corresponds to a set of rays intersecting at a point on the focal plane in the image space. The focal points of an optical system are defined as the pair of points on the principal axis of the system and conjugate to points at infinity.

1.8.2 Principal Points and Principal Planes

- For an optical system, there are two principal points and two principal planes. In order to understand the position of principal points of an optical system, consider a thick convex lens having its principal foci F_1 and F_2 as shown in Fig. 21.
- The rays PQ is incident at point Q on the lens. It is parallel to the principal axis. The ray after refraction through the first surface of lens travels along QR. The ray QR, after refraction through the other surface of a lens travels along RF_2. The point F_2 on the principal axis is called second focal point of the thick convex lens.

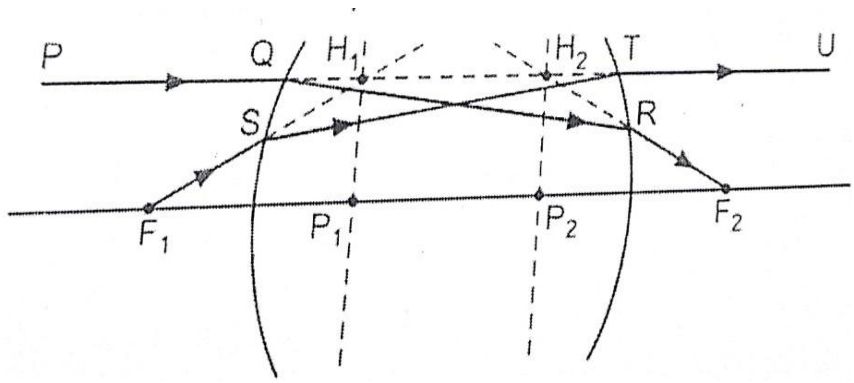


Figure: 21

- If the incident ray PQ and the final emergent ray RF_2 are produced in forward and backward directions respectively, then these rays intersect at point H_2. A plane passing through H_2 and perpendicular to the principal axis is called the second principal plane of the thick convex lens or optical system.

- The point of intersection of this plane with principal axis is called the second principal point (P_2).
- Consider another ray F_1S starting from the first focal point F_1 of the lens. After refraction through the lens, it emerges along the direction TU parallel to the principal axis at the height as that of the ray incident at point Q. The rays F_1S and TU when produced in forward and backward directions respectively, then these rays intersect at point H_1.
- A plane passing through point H_1 and perpendicular to the principal axis is called first principal plane and its point of intersection with the principal axis is called first principal point (P_1).
- In Fig. 21, the incident rays F_1S and PQ are directed towards point H_1 and after refraction these rays appear to come from point H_2, hence H_2 is the image of H_1. The planes H_1P_1 and H_2P_2 are a pair of conjugate planes such that $H_1P_1 = H_2P_2$. The lateral magnification of the planes is +1, so these planes are called the two conjugate planes of unit positive lateral magnification. The distance P_1F_1 is the first principal focal length and P_2F_2 is the second principal focal length of the thick convex lens as shown in Fig. 21.

1.8.3 Nodal Points and Nodal Planes

- Consider an optical system consisting of a thick convex lens as shown in Fig. 22.

- Let F_1 and F_2 be first and second focal points of an optical system respectively and let P_1 and P_2 be first and second principal points respectively. In Fig. 22, H_1P_1 and H_2P_2 are first and second principal planes respectively. AF_1 and BF_2 are first and second focal planes of an optical system respectively.

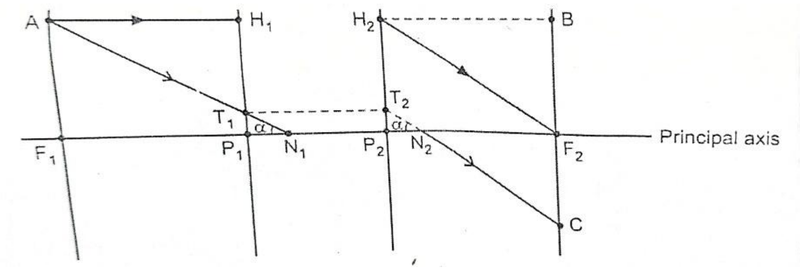


Figure: 22

- Consider any point such as a point A on the first focal plane of an optical system. Consider a ray of light AH_1 travelling parallel to the principal axis. Its conjugate ray emerges along H_2F_2 such that $H_2P_2 = H_1P_1$. Let us take another ray AT_1 parallel to ray H_2F_2 and striking the first principal plane (H_1P_1) at point T_1.
- Its conjugate ray T_2C originates from T_2 such that $T_1P_1 = T_2P_2$. The ray T_2C (or T_2N_2) is parallel to ray H_2F_2 because both the rays originate from the point A taken on the first focal plane of an optical system.
- The points of intersection of incident ray AT_1 and its conjugate emergent rays T_2C with the principal axis are called nodal points. The nodal points N_1 and N_2 are shown in Fig. 22. The planes passing through the nodal points and perpendicular to the principal axis are called the nodal planes. The nodal points N_1 and N_2 are conjugates and any ray like AN_1 and its conjugate emergent ray T_2C are parallel to each other. Nodal points are defined as a pair of conjugate points on the axis having unit positive angular magnification.

- In Fig.22, the triangles $T_1P_1N_1$ and $T_2P_2N_2$ are congruent ($\because \angle T_1P_1N_1 = \angle T_2P_2N_2 = 90^o$, $T_1P_1 = T_2P_2$ and $\angle T_1P_1N_1 = \angle T_2P_2N_2$).

  □ $$P_1N_1 = P_2N_2$$

  Adding the term N_1P_2 on both sides of above equation, we get

  $$P_1N_1 + N_1P_2 = P_2N_2 = N_1P_2$$

  $$P_1P_2 = N_1N_2 \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (8)$$

- Thus, the distance between the principal points P_1 and P_2 is equal to distance between two nodal points N_1 and N_2 of the same optical system.

- In Fig. 22, the triangles AF_1N_1 and $H_2P_2F_2$ are congruent ($\because AF_1 = H_2P_2$, $\angle AN_1F_1 = \angle H_2F_2P_2$ and $\angle AF_1N_1 = \angle H_2P_2F_2$).

  □ $F_1N_1 = P_2F_2$

  But, $F_1N_1 = F_1P_1 + P_1N_1$

  □ $F_1P_1 + P_1N_1 = P_2F_2$

  Or $P_1N_1 = P_2F_2 - P_1F_1$

  But $P_2F_2 = f_2$ and $P_1F_1 = -f_1$

  □ $P_1N_1 = f_2 + f_1$

  But we have shown that $P_1N_1 = P_2N_2$

  □ $P_1N_1 = P_2N_2 = f_1 + f_2$

  If the medium on both sides of the lens is same, then $f_2 = -f_1$.

  □ $P_1N_1 = P_2N_2 = 0$

- From this equation it is clear that, if the medium on both the sides of an optical system is same, the principal points coincide with nodal points.

1.8.4 Cardinal Points for an Optical System of Two Thin Convex Lenses Separated by a Finite Distance

We have studied and obtained an expression for equivalent focal length of the system of two lenses separated by some finite distance (see equation 5) Now, let us find the position of second principal point (P_2) and first principal point (P_1) for a co-axial lens system.

Position of Principal Points:

Position of the second principal point: Let β represent the distance of second principal point of equivalent lens from the second lens L_2. i.e., let $O_2P_2 = \beta$. (Fig.23)

The triangles EF_2O_2 and $H_2F_2P_2$ in Fig. 23 are similar triangles.

□
$$\frac{O_2E}{H_2P_2} = \frac{O_2F_2}{P_2F_2}$$

From Fig. 23 $O_2E = h_2$, $H_2P_2 = h_1$ and $P_2F_2 = f$

□
$$\frac{h_2}{h_1} = \frac{O_2F_2}{f}$$

But, $O_2F_2 = P_2F_2 - P_2O_2 = f - \beta$

□
$$\frac{h_2}{h_1} = \frac{f-\beta}{f}$$

Or
$$\frac{h_1}{h_2} = \frac{f}{f-\beta}$$

But from the Fig. 23, we get

$$\frac{h_1}{h_2} = \frac{f_1}{f_1 - x}$$

□
$$\frac{f}{f-\beta} = \frac{f_1}{f_1 - x}$$

By sign convention, as β is to the left of lens L_2, we take it as negative.

□ $$\frac{f}{f+\beta} = \frac{f_1}{f_1 - x}$$

Or $$ff_1 - xf = ff_1 + \beta f_1$$

□ $$\beta = -\frac{xf}{f_1}$$

- Thus, second principal point P_2 is situated at a distance $\frac{xf}{f_1}$ behind the seconds lens (or to the left of the seconds lens).
- We can also calculate the distance of principal point P_2 from first lens. The distance of the second principal point from the first lens is O_1P_2.

$$O_1P_2 = O_1O_2 - O_2P_2$$

$$= x - \beta$$

$$O_1P_2 = x - \frac{xf}{f_1} = x\left(1 - \frac{f}{f_1}\right)$$

- Thus, second principal point P_2 is situated at a distance $x\left(1 - \frac{f}{f_1}\right)$ in front of the first lens (or to the right of the first lens).
- Since H_2P_2 is the position of equivalent lens, it is clear that equivalent lens must be placed at a distance $\frac{xf}{f_1}$ to the left of the second lens or $x\left(1 - \frac{f}{f_1}\right)$ to the right of the first lens.

Position of the First Principal Point:

- To find the position of first principal point, consider Fig. 24.
- In figure, let F_1E be incident ray on lens L_1 at a height h_1 above the principal axis. This ray after refraction through the first lens travels along path EQ. The ray EQ is again incident on lens L_2 at height h_2 above the principal axis.

- The ray EQ after refraction through lens L_2 travels along QP. The final emergent ray QP is parallel to principal axis. The final emergent ray QP and the incident ray F_1E, when produced intersect at point H_1. The plane passing through H_1 and perpendicular to the principal axis is called as the first principal plane and it is shown as H_1P_1 in Fig. 24. Its point of intersection with the principal axis is called first principal point (P_1).

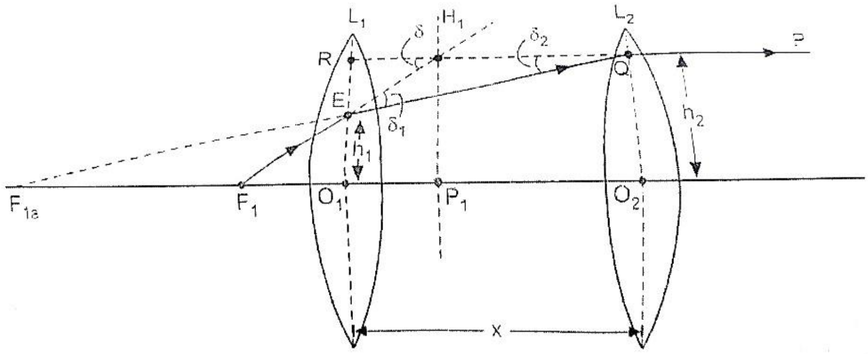


Figure: 24

- The ray EQ when produced in backward direction, intersect the principal axis at point F_{1a}. As the emergent ray QP is parallel to the principal axis, the point F_{1a} is the first principal focal point of the second lens. Let the distance of first principal point P_1 from the first lens be α, i.e. $O_1P_1 = \alpha$.

The triangles EF_1O_1 and $H_1F_1P_1$ in Fig. 24 are similar triangles.

□ $$\frac{O_1E}{H_1P_1} = \frac{O_1F_1}{P_1F_1}$$

But $O_1E = h_1$, $H_1P_1 = h_2$ and $P_1F_1 = f$

□ $$\frac{h_1}{h_2} = \frac{O_1F_1}{f}$$

From Fig. 24, $O_1F_1 = P_1F_1 - O_1P_1 = f - \alpha$

□ $$\frac{h_1}{h_2} = \frac{f-\alpha}{f} \qquad \ldots\ldots\ldots\ldots\ldots (9)$$

- The point P_1 is to the right of the first lens, hence according to sign conventions, O_1P_1 is positive.

In Fig. 24, $h_1 = O_1E = O_1R - ER$

$$= h_2 - x\,\delta_2$$

$$= h_2 - x\,\frac{h_2}{f_2}$$

 $$\frac{h_2}{f_2} = \left(1 - \frac{x}{f_2}\right) \qquad \text{...... (10)}$$

Equating equations (9) and (10), we get

$$\frac{f - \alpha}{f} = 1 - \frac{x}{f_2}$$

$$1 - \frac{\alpha}{f} = 1 - \frac{x}{f_2}$$

$$\frac{\alpha}{f} = \frac{x}{f_2}$$

 $$\alpha = \frac{xf}{f_2} \qquad \text{........... (11)}$$

- Thus, the first principal point P_1 of the equivalent lens is situated at a distance $\frac{xf}{f_2}$ from the first lens is measured to the right of lens L_1.
- In Fig. 25, F_1 is the first focal point of the equivalent lens and is situated at a distance f towards the left of the principal point P_1. F_2 is the second focal point of the equivalent lens and is situated at a distance f towards the right of the second principal point P_2.

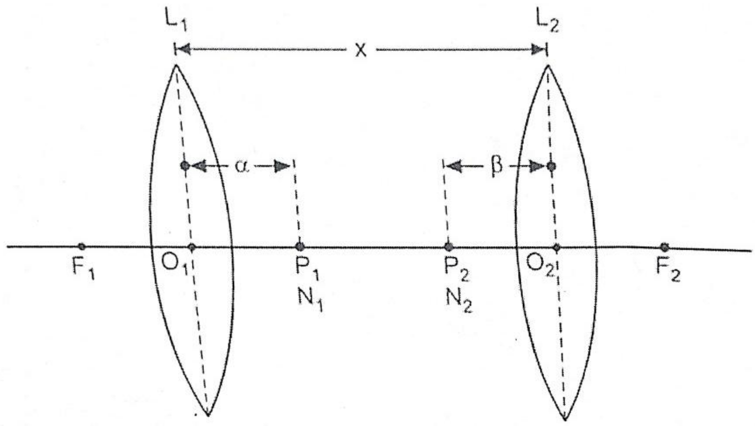


Figure: 25

- Since the medium on both sides of an optical system is same, the principal points coincide with nodal points as shown in Fig. 25.

1.9 GENERAL FORMULATION FOR THE LOCATION OF CARDINAL POINTS :

The imaging process for formation of image of an object by an optical system may be described as a combination of the following processes :

(i) Translation in object space represented by matrix $[T_1]$

(ii) Refraction in medium represented by matrix [S]

and (iii) Translation in image space represented by matrix $[T_2]$

If $\begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix}$ is column vector representing image and $\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix}$ is the column vector representing the image. Then

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = [T_2][S][T_1]\begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \qquad \text{……………………. (12)}$$

$\boxed{\text{The refraction matrix, In general, is expressed as}}$

$$[S] = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \qquad \text{……….(13)}$$

This matrix is also called optical element matrix.

The resultant of metrix [T_2] [S] [T_1] is called the overall system matrix denoted by N i.e.,

$$N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} = [T_2][R][T_1]$$

If d_1 is the object distance and d_2 the image distance, then by convention d_1 is positive and d_2 is negative.

So, $$[T_1] = \begin{bmatrix} 1 & +d_1 \\ 0 & 1 \end{bmatrix}$$

$$[T_2] = \begin{bmatrix} 1 & -d_2 \\ 0 & 1 \end{bmatrix}$$

Thus, $$N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} = \begin{bmatrix} 1 & -d_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 1 & +d_2 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -d_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_{11} & S_{11}d_1 & + & S_{12} \\ S_{21} & S_{21}d_2 & + & S_{22} \end{bmatrix}$$

$$= \begin{bmatrix} S_{11} - d_2 S_{21} & S_{11}d_1 + S_{12} - S_{21}d_1 d_2 - S_{22}d_2 \\ S_{21} & S_{21}d_1 + d_{22} \end{bmatrix}$$

$$\Rightarrow \left.\begin{aligned} N_1 &= S_{11} - S_{21}d_2 \\ N_2 &= S_{11}d_1 + S_{12} - d_2 - (S_{22} + S_{21}d_1 \\ N_{22} &= S_{21} \end{aligned}\right\} \text{.......}(14)$$

$$N_{21} = S_{21}d_1 + S_{22}$$

$$\therefore \quad \begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \quad \text{...................}(16)$$

$$\Rightarrow \quad y_2 = N_{11}y_1 + N_{12}\alpha_1 \quad \text{.........}(17)$$

and $$\alpha_2 = N_{21}y_1 + N_{22}\alpha_1 \quad \text{..........}(18)$$

The four elements of overall system matrix may be used to locate the positions of the cardinal points of the system.

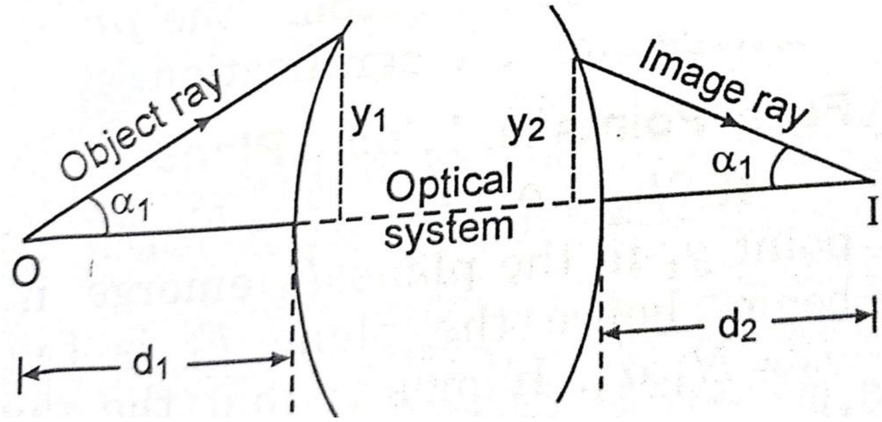


Figure: 26

Principal Points and Principal Planes:

It $N_{12} = 0$, $y_2 = N_1 y_1$. It means that all the rays from the plane H_1 containing the object y_1 will pass through a point in the plane H_2 containing the image point y_2, thus the plane H_1 and H_2 are conjugate planes. If these planes also satisfy the condition $m_y = \frac{y_2}{y_1} = +1,$ then these **conjugate planes become the principal planes or the planes of unit positive lateral magnification.**

Focal Points and Focal Planes:

If $N_{22} = 0$, then $\alpha_2 = N_{21} y_1$. It means that the rays from the object point y_1 in the plane F_1 emerge in the same direction i.e., as a parallel beam, hence F_1 is the first focal plane. If $N_{11} = 0$, then $y_2 = N_{12}\alpha_1$. It means that the rays incident as a parallel beam pass through a point y_2, thus the plane passing through y_2 is the second focal plane.

Nodal Points and Nodal Planes:

If $N_{22} = 0$, $\alpha_2 = N_{22}\alpha_1$. This implies that emergent ray direction is determined by incident ray direction. Thus, if a collimated beam enters the system, the emergent beam is also collimated. Such a system is said to be collimated. In addition, if $\frac{\alpha_2}{\alpha_1} = N_{22} = 1$ the conjugate points on the principal axis are the nodal points and the planes passing through them are the nodal planes.

1.10 CARDINAL POINTS OF TWO THIN LENSES SEPARATED BY A DISTANCE

Consider two thin convex lenses of focal lengths f_1 and f_2, separated by a distance d. The refraction matrices for the two lenses are $[R_1]$ and $[R_2]$;

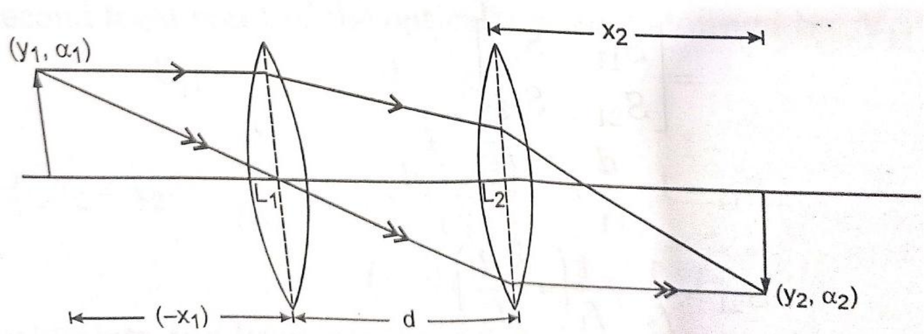


Figure: 27

while the translation matrix between the lenses is [T]. If object ray is represented by $\begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix}$ image ray by $\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix}$, x_1 is object distance from first lens, x_2 is image distance from second lens, then by matrix method

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & -x_2 \\ 0 & 1 \end{bmatrix} [R_2][T][R_1] \begin{bmatrix} 1 & x_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \qquad \ldots\ldots\ldots\ldots\ldots(18)$$

where the matrix $\begin{bmatrix} 1 & x_1 \\ 0 & 1 \end{bmatrix}$ represents translation matrix from object to the first lens and $\begin{bmatrix} 1 & x_2 \\ 0 & 1 \end{bmatrix}$ represents translation matrix from second lens to image.

The magnitudes of x_2 and x_2, are negative since both rays are converging (fig. 27). But the sign convention x_1 is negative and x_2 is positive

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & -x_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{f_2} & 1 \end{bmatrix} \begin{bmatrix} 0 & -d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{f_2} & 1 \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \qquad (19)$$

The refraction matrix for lens system is

$$R = \begin{bmatrix} 1 & 0 \\ \frac{1}{f_2} & 1 \end{bmatrix} \begin{bmatrix} 0 & -d \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{f_1} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -d \\ \frac{1}{f_2} & -\frac{d}{f_2} + 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{1}{f_2} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 - \frac{d}{f_1} & -d \\ \frac{1}{f_2} + \frac{1}{f_1}\left(1 - \frac{d}{f_2}\right) & 1 - \frac{d}{f_2} \end{bmatrix} \tag{20}$$

$$= \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \tag{21}$$

where $S_{11} = 1 - \frac{d}{f_1}$; $S_{12} = -d$

$$S_{21} = \frac{1}{f_2} + \frac{1}{f_1}\left(1 - \frac{d}{f_2}\right)$$

$$f_{22} = 1 - \frac{d}{f_2} \tag{22}$$

$$\therefore \quad \begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & -x_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{y_1}{\alpha_1} \end{bmatrix}$$

$$= \begin{bmatrix} S_{11} - S_{21}x_2 & S_{12} - S_{22}x_2 \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix}$$

$$= \begin{bmatrix} S_{11} - S_{21}x_2 & x_1(S_{11} - S_{21}x_2) + S_{12} - S_{22}x_2 \\ S_{21} & x_1 S_{21} + S_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \tag{23}$$

If N is overall system matrix, then

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = [N] \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \tag{24}$$

with $[N] = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$

where $$\left.\begin{aligned} N_{11} &= S_{11} - S_{21}x_2 \\ N_{12} &= x_1(S_{11} - S_{21}x_2) + S_{12} - S_{22}x_2 \\ N_{21} &= S_{21} \end{aligned}\right] \tag{25}$$

and $$N_{22} = x_1 S_{21} + S_{22}$$

1.10.1 Focal Points and Focal Planes

The first focus of the system is defined by the system matrix element $N_{22} = 0$

$$\Rightarrow \quad x_1 S_{21} + S_{22} = 0$$

$$\Rightarrow \quad x_1 = -\frac{S_{22}}{S_{21}} = -\frac{1-\frac{d}{f_2}}{\frac{1}{f_2}+\frac{1}{f_1}\left(1-\frac{d}{f_2}\right)} = -\frac{1-\frac{d}{f_2}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2}} \tag{26}$$

$$\therefore \quad L_1 F_1 = -\frac{1-\frac{d}{f_2}}{\frac{1}{f_2}+\frac{1}{f_2}-\frac{d}{f_1 f_2}}$$

The second focal point of the optical system is defined by $N_{11} = 0$

$$\therefore \quad S_{11} - S_{21}x_2 = 0$$

$$\Rightarrow \quad L_2 F_2 = x_2 = +\frac{S_{11}}{S_{21}} = +\frac{1-\frac{d}{f_1}}{\frac{1}{f_2}+\frac{1}{f_1}\left(1-\frac{d}{f_2}\right)} = +\frac{1-\frac{d}{f_1}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2}} \tag{27}$$

1.10.2 Principal Points and Principal Planes

The principal planes of optical system are defined by $N_{12} = 0$ and $N_{11} = 1$

$N_{12} = 0$ gives $-x_1(S_{11} - S_{21}x_2) + S_{12} - S_{22}x_2 = 0$

and $N_{11} = 1$ gives $S_{11} - S_{21}x_2 = 1$

$$\Rightarrow \quad x_2 = \frac{S_{11}-1}{S_{21}} = -\frac{\frac{d}{f_1}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2}} \tag{28}$$

The position of second principal point (H_2) from second lens (L_2) is

$$L_2H_2(= x_2) = -\frac{\frac{d}{f_1}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}} \tag{29}$$

Also $\quad -x_1(S_{11} - S_{21}x_2) + S_{12} - S_{22}x_2 = 0$

As $\quad S_{11} - S_{21}x_2 = 1 \quad \therefore \quad -x_1 S_{12} - S_{22}x_2 = 0$

$$\Rightarrow \quad x_1 = S_{12} - S_{22}x_2 = -\frac{\frac{d}{f_2}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}}$$

$$\therefore \quad L_1H_1 = x_1 = -\frac{\frac{d}{f_1}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}} \tag{30}$$

The two focal points (F_1), (F_2) and two principal points (H_1, H_2) are shown in fig.

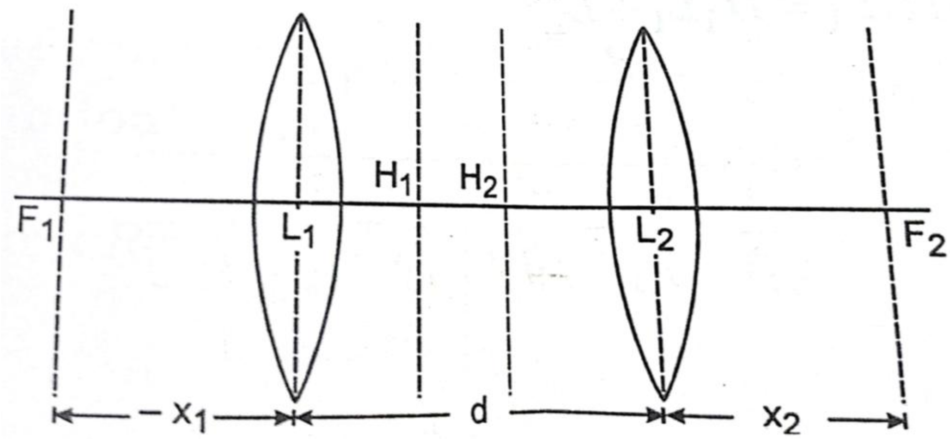


Figure: 28

1.10.3 Significance of Principal Planes

As already pointed out, principal planes are the conjugate planes having unit positive transverse magnifications. It means that if a ray of light is incident on first principal plane at some height, the corresponding emergent ray leaves the second principal plane at the same height.

Consider two thin convex lenses L_1 and L_2 of focal lengths f_1 and f_2 separated by a distance d.

A light ray PA coming parallel to principal axis, strikes the first lens at A, it follows the path AB and strikes the second lens at B and finally follows the path BF_2. F_2 is the second focal point. If incident ray PA and emergent ray BF_2 are produced they meet at A_2 and perpendicular A_2H_2 dropped from A_2 on the principal axis is the second principal plane. The equivalent single lens may be placed at H_2 having focal length H_2F_2. Thus, the distance H_2F_2 is the second focal length of optical system.

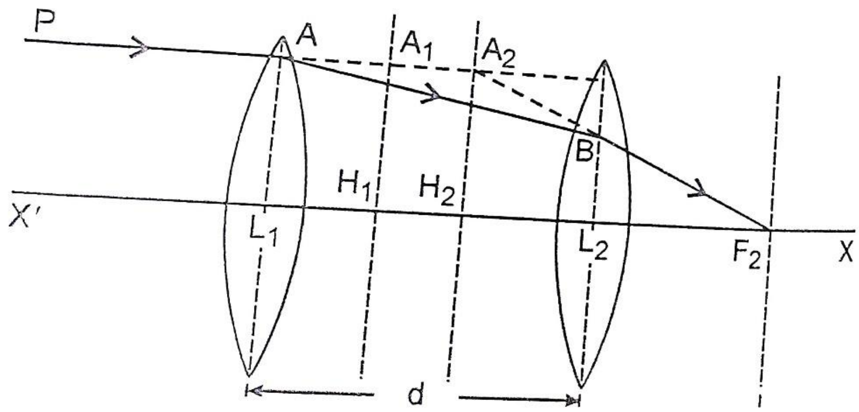


Figure: 29

The equivalent focal length of optical system

$H_2F_2 = |H_2L_2| + |L_2F_2|$

$$= +\frac{\frac{d}{f_1}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}} + \frac{1-\left(\frac{d}{f_1}\right)}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}}$$

If F_e is the equivalent focal length of lens system, then $H_2F_2 – F_e$, so

$$F_e = \frac{1}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}} \Rightarrow \frac{1}{F_e} = \frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1f_2}$$

So, $L_2H_2 = -\frac{d\,F_e}{f_1}$

If F_1 is the first focal point of optical system, then

$H_1F_1 = H_1L_1 + L_1F_1$

$$= \frac{\frac{d}{f_2}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2}} + \frac{1-\frac{d}{f_2}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2}}$$

$$= \frac{1}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{d}{f_1 f_2}}$$

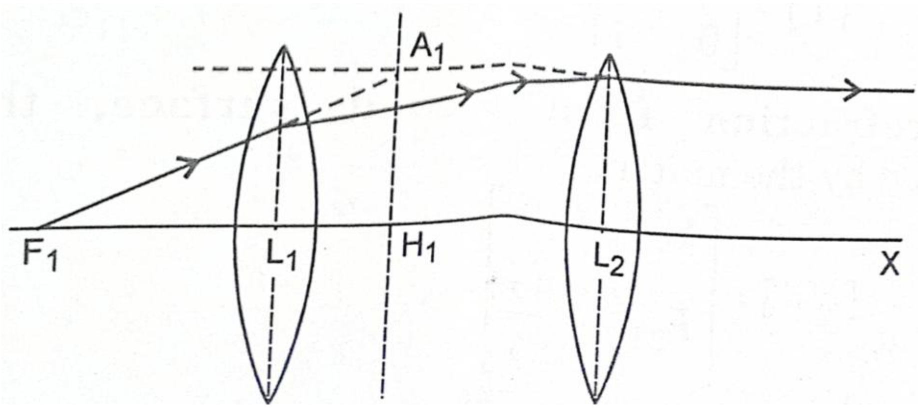


Figure: 30

But H_1F_1 = first focal length of optical system

$$= -Fe'$$

$$-\frac{1}{F_e'} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{d}{f_1 f_2}$$

Obviously

$$\boxed{F_e = -F_e'}$$

That is two focal lengths of optical system are equal and opposite in sign.

1.11 APLANATIC POINTS OF A SPHERICALLY REFRACTING SURFACE AND A SPHERE

A surface is aplanatic with respect to two conjugate points if all the rays originating from one of the two points, after refraction or reflection at the surface, converge to or appear to diverge from another points. This property of the surface is called aplanatism. Such a pair of points are termed as aplanatic foci or aplanatic points.

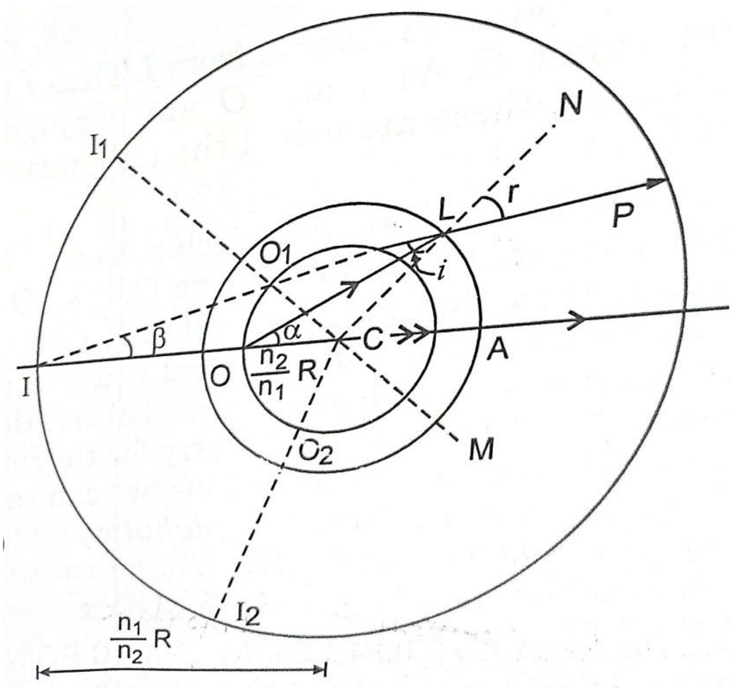


Figure: 31

Let LAM be a section of a sphere of radius R having centre C. Let the refractive index of the medium of the sphere be n_1 and that of the medium outside the sphere be n_2 where $n_1 > n_2$. Let O be a point object on the axis of the surface LAM at a distance $\frac{n_2}{n_1}$ R from the centre of the sphere. A ray OL incident on the surface at point L, after refraction through the surface, is bent away from the normal CLN and follows the path LP. Another ray OA, coincident with the axis of LAM and incident on the surface normally at A, undergoes no deviation and passes straight. Both the refracted rays LP and OA, when produced, intersect at I. Thus, I is the virtual image of O.

Let $\angle LOC = \alpha$, $\angle OIL = \beta$. If i and r are the angles of incidence and refraction, we have

$$\angle OLC = i,\ \angle PLN = \angle ILC = r.$$

In ΔLOC, $\quad \frac{\sin i}{\sin \alpha} = \frac{OC}{LC}.$

But $CL = R, OC = \frac{n_2 R}{n_1}$

$\therefore \quad \frac{\sin i}{\sin \alpha} = \frac{n_2 R/n_1}{R}$

$= \frac{n_2}{n_1}$ (31)

According to Snell's law

$\frac{\sin i}{\sin \alpha} = \frac{n_2}{n_1}$ (32)

$\therefore$ From equations (31) and (32).

$\frac{\sin i}{\sin \alpha} = \frac{\sin i}{\sin r}$

$\therefore \quad \sin \alpha = \sin r$

or $\alpha = r$ (33)

In Δ ILO, exterior $\angle$LOC = β +$\angle$ILO

or $\alpha = \beta + (r - i)$ (sin $\angle$ILO = r – i)

or $r = \beta + r - i$ [from (33)]

$\therefore$ $\beta = i$, (34)

In Δ ILC, $\frac{\sin \beta}{\sin r} = \frac{CL}{IC}$;

$\frac{\sin i}{\sin r} = \frac{R}{IC}$; ($\because$ β = i and CL = R)

$CI = R \frac{\sin r}{\sin i} = R.\frac{n_1}{n_2}$ $\left[\because \frac{\sin r}{\sin i} = \frac{n_1}{n_2} \; from \; (2)\right]$

$$CI = \frac{n_1 R}{n_2}$$

Therefore, if the point object is placed on the axis of the refracting surface at distance $OC = \frac{n_2 R}{n_1} = 1n_2 R$ R from the centre of the sphere, the corresponding image is formed at a distance $CI = \frac{n_1 R}{n_2} = \frac{R}{1n_2}$ from the centre C towards the object. Similarly, if the object is placed at a distance $\frac{n_1 R}{n_2}$ from the centre, the corresponding image is formed at a distance $\frac{n_2 R}{n_1}$ from the centere. This result is independent of position of the L, i.e., the angle α which the incident ray makes with the principal axis. Therefore, all the rays, whether paraxial or marginal, starting from O are refracted such that they appear to diverge from I. Thus, I is the perfect point image of the point object O. As I and O are conjugate points free from spherical aberration, these are called the aplanatic points of the spherical surface LAM.

Now let us draw two circles around C with radii CI and CO. If we choose a series of point objects O_1, O_2, O_3 …. etc. on any one of the two circles, then the images of all these point objects I_1, I_2, I_3 ….. etc., after refraction through the spherical refracting surface, will be formed on the second surface free from spherical aberration, because any line through the centre C is the axis of symmetry for the image formation. If we generalize these ideas to three dimensions, we can say that a complete sphere has an infinite number of pairs of aplanatic points situated on a pair of conjugate surfaces O_1OO_2, and I_1II_2. These two surfaces are called the aplanatic surface of the sphere.

1.11.1 Importance of Aplanatic Points

(i) **Aplanatic lens:** Aplanatic lens which is free from the defects of spherical aberration and coma is called an aplanatic lens. The pair of conjugate points free from spherical aberration and coma are called aplanatic points. According to the property of aplanatic points if a point object is placed on an aplanatic point of the spherical refracting surface, its image free from spherical aberration and coma will be formed on the other aplanatic point. For a single spherical surface these points lie at distance nR and R/n from the centre of curvature. An aplanatic lens is mostly used as the front lens of a high-power microscope objective. This is due to the fact that the objective of a high-power microscope receives a wide-angled pencil of rays diverging from the object.

(ii) **Oil immersion lens :** It consists of a hemispherical lens of radius R with is plane surface closer to the object. Its plane surface is in contact with the cedar wood oil of height R/n, where n is the refractive index of the lens (i.e., refractive index of lens = refractive index of cedar wood oil). O is a point object whose image free from spherical aberration is formed at I (fig. 32). From, the property of aplanatic points, the point I will be at a distance nR from C, i.e.,

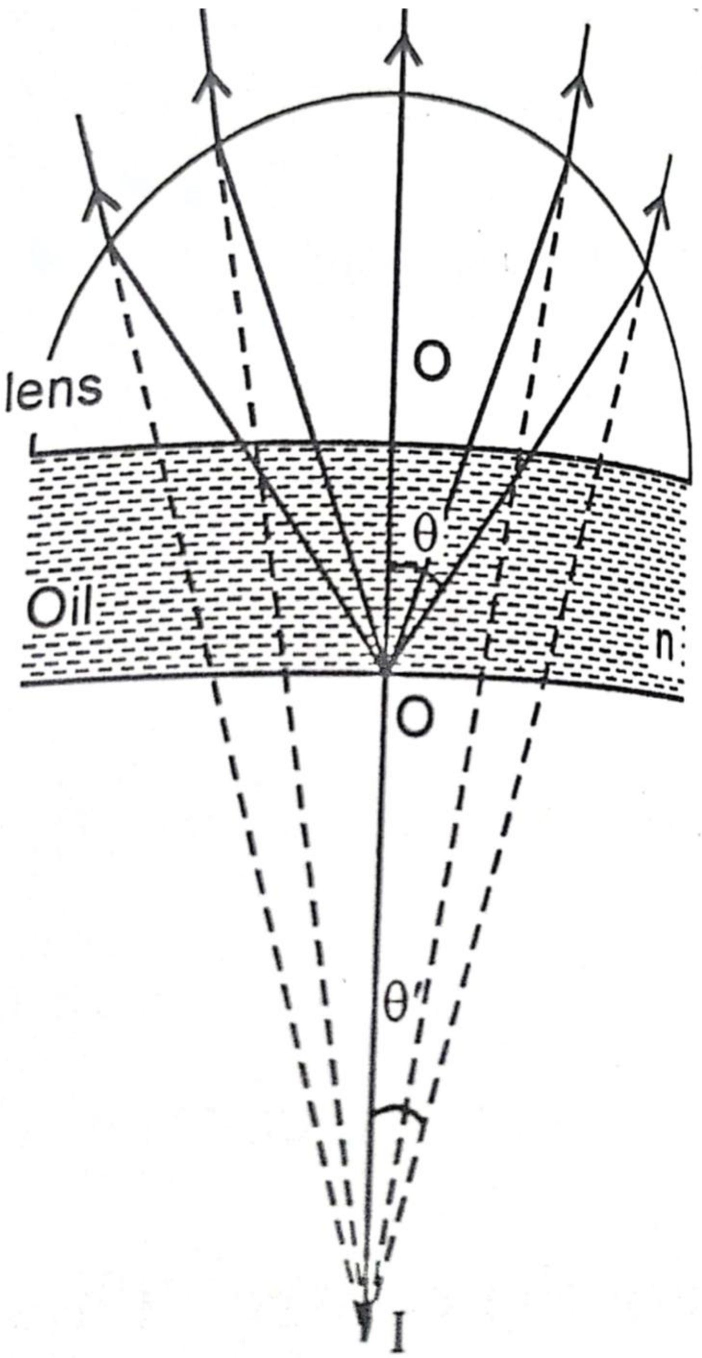


Figure: 32

CI = nR

The semi-angle θ of the cone of incident rays from O is given by

$$\tan\theta = \frac{R}{R/n} = n$$

while the semi-angle θ' if the emergent cone of rays is given by

$$\tan\theta' = \frac{R}{nR} = 1/n$$

(This lens is used in the objective of the microscope.)

1.12 SYSTEM MATRIX AND CARDINAL POINTS FOR A THICK LENS

Consider a lens of thickness t and refractive index n_2, placed in media of refractive indices n_1 and n_3. Let R_1 and R_2 be the radii of curvature of the two surfaces of the lens.

Suppose light ray OA is incident on the lens at height y_1 making angle α_1 with the principal axis. The path of ray inside the lens is AB and final image ray is BI. The column vectors of object ray is $\begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix}$ and that of image ray is $\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix}$.

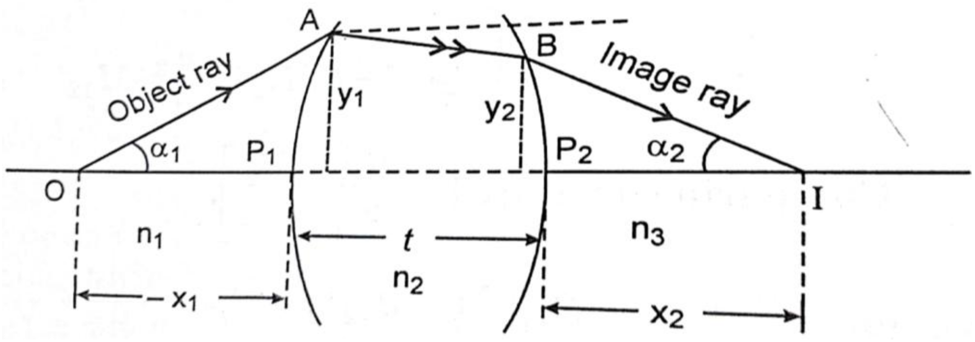


Figure: 23

The image forming process undergoes the following three operations:

(i) **Refraction at first surface;** the refraction process is defined by matrix

$R_1' = \begin{bmatrix} 1 & 0 \\ M_{12} & \frac{n_1}{n_2} \end{bmatrix}$,where $M_{12} = \frac{n_2 - n_1}{n_2 R}$

(ii) **The translation from P1 to P2,** the translation process is defined by matrix [T].

$[T] = \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix}$

(iii) **The refraction from second surface,** the refraction process is defined by the matrix

$$[R_2'] = \begin{bmatrix} 1 & 0 \\ P_{23} & \frac{n_2}{n_3} \end{bmatrix}$$

Thus, image forming process is given by

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = [R_2'][T][R_1']\begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \tag{35}$$

The system matrix can be written as

$$[S] = [R_2'][T][R_1']$$

$$= \begin{bmatrix} 1 & 0 \\ P_{23} & \dfrac{n_2}{n_3} \end{bmatrix}\begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ M_{12} & \dfrac{n_1}{n_2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ P_{23} & \dfrac{n_2}{n_3} \end{bmatrix}\begin{bmatrix} 1 - t.M_{12} & -\dfrac{n_1}{n_2}t \\ M_{12} & \dfrac{n_1}{n_2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 - t\,M_{12} & -\frac{n_1}{n_2}t \\ P_{23}(1 - t\,M_{12}) + \frac{n_2}{n_3}\,M_{12} & -\frac{n_1}{n_2}t\,P_{23} + \frac{n_1}{n_3} \end{bmatrix} \tag{36}$$

Comparing it with $[S] = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & 22 \end{bmatrix}$, (37)

we get $S_{11} = 1 - M_{12}t$; $S_{12} = -\frac{n_1}{n_2}t$

$S_{21} = P_{23}(1 - t\,M_{12}) + \frac{n_2}{n_1}M_{12}$

and $S_{22} = -\frac{n_1}{n_2}t\,P_{23} + \frac{n_1}{n_3}$

Now $M_{12} = \frac{n_2 - n_1}{n_2 R_1} = \frac{n_1}{n_2 f_1}$ where $\frac{1}{f_1} = \frac{n_2 - n_1}{n_2 R_1}$

and $P_{23} = -\frac{n_3 - n_2}{n_3 R_2} = \frac{1}{f_2}$

$$\left.\begin{array}{c} S_{11} = 1 - \frac{n_1 t}{n_2 f_1}, S_{12} = -\frac{n_1}{n_2} t \\ S_{21} = \frac{1}{f_2}\left(1 - \frac{n_1 t}{n_2 f_1}\right) + \frac{1}{f_1} \\ S_{22} = -\frac{n_1 t}{n_2 f_2} + \frac{n_1}{n_3} \end{array}\right\} \quad (38)$$

For overall image formation

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 1 & -x_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -x_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} S_{11} & x_1 S_{11} + S_{12} \\ S_{21} & x_1 S_{21} + S_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix}$$

$$= \begin{bmatrix} S_{11} - x_2 S_{21} & x_1 S_{11} + S_{12} - x_2(x_1 S_{21} + S_{22}) \\ S_{21} & x_1 S_{21} + S_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \quad (39)$$

If N is overall system matrix, then

$$\begin{bmatrix} y_2 \\ \alpha_2 \end{bmatrix} = [N] \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ \alpha_1 \end{bmatrix} \quad (40)$$

Comparing (39) and (40), we get

$$\left.\begin{array}{c} N_{11} = S_{11} - x_2 S_{21}, N_{12} = x_1 S_{11} + S_{12} - x_2(x_1 S_{21} + S_{22}) \\ N_{21} = S_{21} \; and \; N_{22} = x_1 \, S_{21} + S_{22} \end{array}\right\} \quad (41)$$

so

and $$\left.\begin{array}{c} y_2 = N_{11} y_1 + N_{12} \alpha_1 \\ \alpha_2 = N_{21} y_1 + N_{22} \alpha_1 \end{array}\right\} \quad (42)$$

Cardinal Points:

First Principal Focus: First focal point is a point on the principal axis for which the conjugate image point is at infinity. i.e., ($x_2 = \infty$). In equation (42), if we put $N_{22} = 0$, we get $\alpha_2 = N_{21}\, y_1$.

This means that the rays passing through the object point for any value of α_1, after emergence from the lens have same value of α_2 i.e., the rays become parallel. This is the characteristic of first principal focal plane. Thus, the first principal focal plane is defined by the transformation matrix element $N_{22} = 0$

$\Rightarrow \qquad x_1 S_{21} + S_{22} = 0 \qquad \Rightarrow \qquad x_1 = -\frac{S_{22}}{S_{21}}$

Thus $\qquad P_1F_1 = x_1 = \frac{\frac{n_1}{n_2}\frac{t}{f_2} - \frac{n_1}{n_3}}{\frac{1}{f_2}\left(1 - \frac{n_1}{n_2}\frac{t}{f_1}\right) + \frac{1}{f_1}}$

Distance of first principal focus from pole of first surface

$$\Rightarrow \qquad P_1F_1 = -\frac{\frac{n_1}{n_3}\left(1 - \frac{n_3}{n_2}\frac{t}{f_2}\right)}{\frac{1}{f_1} + \frac{1}{f_2} - \frac{n_1 t}{n_2 f_1 f_2}} \qquad (43)$$

Second Principal Focus: The second focus is the point on the principal axis where all the rays parallel to principal axis converge at or appear to come from this point

If we put $N_{11} = 0$ in equation (42), we get $y_2 = N_{11}\alpha_1$. This means that all the incident rays having same value of α_1 (i.e., incident parallel rays) pass through the same image point y_2. Thus, the plane passing through y_2 is the second focal plane and is point of intersection with the principal axis it the second focal point.

Now putting $N_{11} = 0$ in equation (41), we get

$$S_{11} - x_2 S_{21} = 0$$

$$\Rightarrow \qquad x_2 = \frac{S_{11}}{S_{21}}$$

Using (38), we get distance of second focal point from second surface

$$P_2F_2 = x_2 = \frac{S_{11}}{S_{21}} = \frac{1 - \frac{n_1 t}{n_2 f_1}}{\frac{1}{f_2}\left(1 - \frac{n_1}{n_2}\frac{t}{f_1}\right) + \frac{1}{f_1}}$$

$$\Rightarrow \qquad P_2F_2 = \frac{1 - \frac{n_1 t}{n_2 f_1}}{\frac{1}{f_1} + \frac{1}{f_2} - \frac{n_1 t}{n_2 f_1 f_2}} \qquad (44)$$

Principal Points and Principal Planes:

The principal planes are the conjugate planes having unit positive lateral magnification. If we set $N_{12} = 0$, in eqn. (42), we get

$$y_2 = N_{11} y_1 \;\Rightarrow\; \frac{y_2}{y_1} = N_{11}$$

Therefore, for unit magnification $\frac{y_2}{y_1} = 1$

This implies that the conditions $N_{11} = 1$ and $N_{12} = 0$ define completely the principal planes.

Now putting $N_{11} = 1$ and $N_{12} = 0$ in eqn. (39), we get

$$S_{11} - x_2 S_{21} = 1 \;\Rightarrow\; x_2 = \frac{S_{11}-1}{S_{21}} \qquad (45)$$

and $\qquad x_1 S_{11} + S_{12} - x_2(x_1 S_{21} + S_{22}) = 0$

$\Rightarrow \qquad x_1 S_{11} + S_{12} - \frac{S_{11}-1}{S_{21}}(x_1 S_{21} + S_{22}) = 0$

$\Rightarrow \qquad x_1\left(S_{11} - \frac{S_{11}-1}{S_{21}}.S_{21}\right) + S_{12} - \frac{(S_{11}-1)S_{22}}{S_{21}} = 0$

$\Rightarrow \qquad x_1 = \frac{(S_{11}-1)S_{22}}{S_{21}} - S_{22} \qquad (46)$

Thus, the distance of second principal plane from second surface is

$$P_2F_2(= x_2) = \frac{S_{11}-1}{S_{21}} = \frac{\left(1 - \frac{n_1}{n_2}\frac{t}{f_1}\right) - 1}{f_2\left(1 - \frac{n_1 t}{n_2 f_1}\right) + \frac{1}{f_1}}$$

$$= \frac{\frac{n_1}{n_2}.\frac{t}{f_1}}{\frac{1}{f_1} + \frac{1}{f_2} - \frac{n_1 t}{n_2 f_1 f_2}} \qquad (47)$$

and the distance of first principal plane from first surface is

$$P_1H_1(= x_2) = \frac{(S_{11}-1)S_{22}}{S_{21}} - S_{12}$$

Using eqn. (42), we get

$$P_1H_1 = \frac{\left(1 - \frac{n_1}{n_2}\frac{t}{f_1} - 1\right).\left(-\frac{n_1}{n_2}\frac{t}{f_2} + \frac{n_1}{n_3}\right)}{\frac{1}{f_2}\left(1 - \frac{n_1 t}{n_2 f_2}\right) + \frac{1}{f_1}} - \frac{n_1}{n_2}t$$

$$= \frac{\frac{n_1}{n_2}\frac{t}{f_1}\left(\frac{n_1}{n_2}\frac{t}{f_2} \; - \frac{n_1}{n_3}\right)}{\frac{1}{f_1} + \frac{1}{f_2} - \frac{n_1 t}{n_2 f_1 f_2}} - \frac{n_1}{n_2} t$$

Nodal Points and Nodal Planes:

The nodal planes are conjugate planes having unit angular magnification. Therefore, if we put $N_{21} = 0$ in equation (42, we get

$$\alpha_2 = N_{22}\alpha_1$$

$\Rightarrow$ Angular magnification $\frac{\alpha_2}{\alpha_1} = N_{22}$

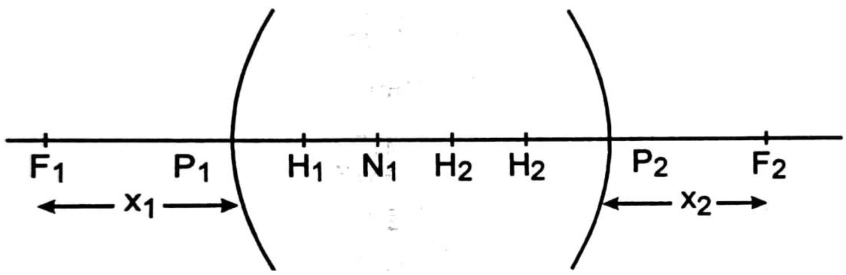
This implies that the nodal planes are characterized by

$N_{21} = 0$ and $N_{22} = 1$

In view of these conditions, equations (41) gives

$S_{12} = 0$ and $x_1 S_{21} + S_{22} = 1$

Distance of first nodal plane from first surface is



$$x_1 = \frac{1-S_{22}}{S_{21}} = \frac{1+\frac{n_1}{n_2}\frac{t}{f_2}-\frac{n_1}{n_3}}{\frac{1}{f_2}\left(1-\frac{n_1 t}{n_2 f_1 f_2}\right)+\frac{1}{f_1}} \qquad \therefore \qquad P_1N_1 = \frac{1+\frac{n_1}{n_3}+\frac{n_1}{n_2}\frac{t}{f_2}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{n_1 t}{n_2 f_1 f_2}}$$

Also, $P_2N_2 = P_2H_2 - H_2N_2$

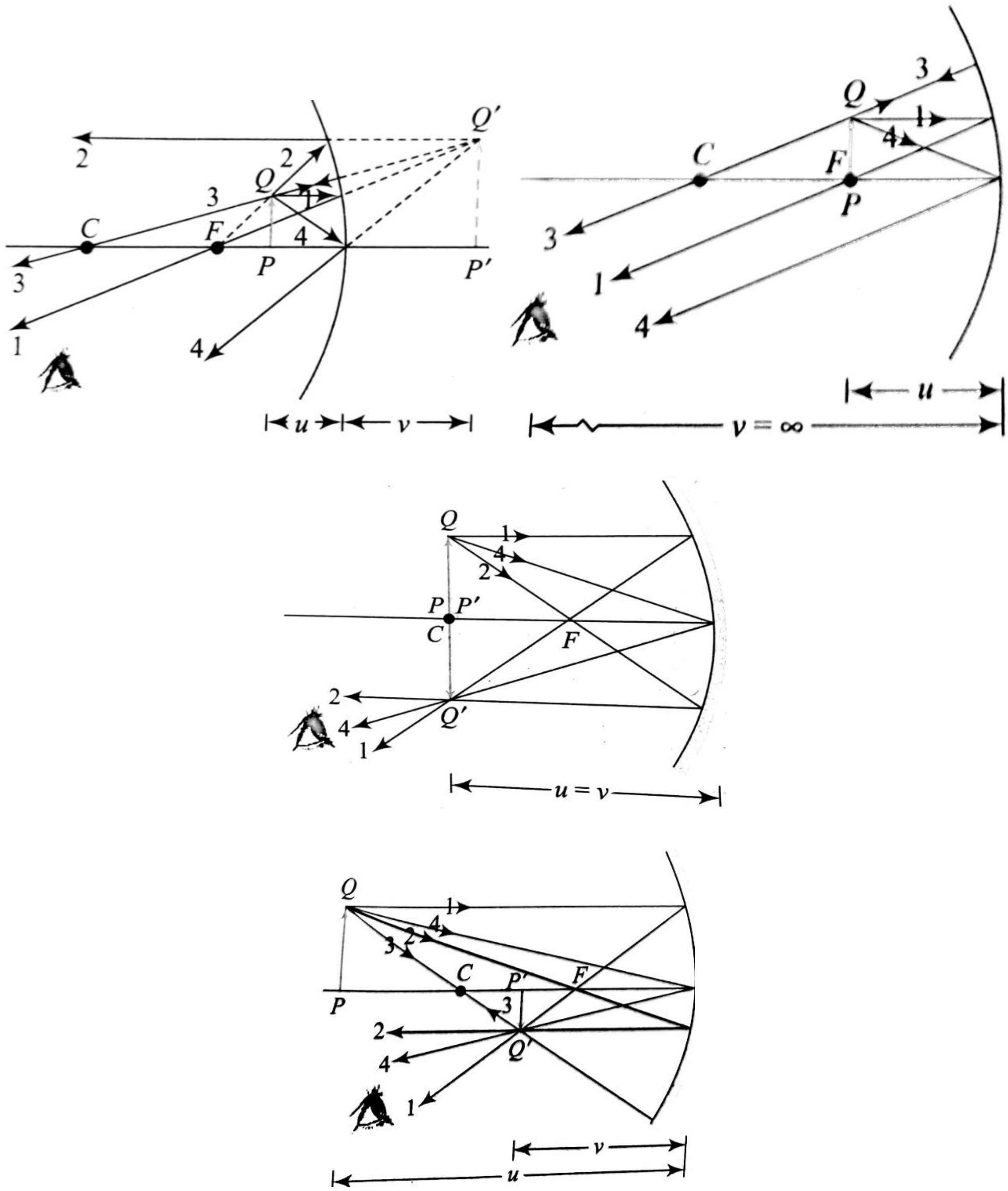
As $H_1N_1 = H_2N_2$

$\therefore$ $P_2N_2 = |P_2H_2| - |H_1N_1|$

$$= \left[\frac{1-\frac{n_3}{n_1}+\frac{n_3}{n_2}\frac{t}{f_1}}{\frac{1}{f_1}+\frac{1}{f_2}-\frac{n_1 t}{n_2 f_1 f_2}}\right]$$

1.13 RULES FOR RAY DIAGRAMS

The position size, and nature of the images formed by mirrors are conventionally expressed by ray diagrams (see figure).



We can locate the image of any extended object graphically by drawing any two of the following four special rays:

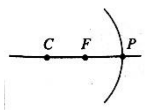
(a) A ray initially parallel to the principal axis is reflected through the focus of the mirror (1).

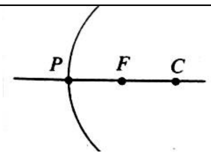
(2) A ray passing through the centre of curvature is reflected back along itself (3).

(c) A ray initially passing through the focus is reflected parallel to the principal axis (2).

(d) A ray incident at the pole is reflected symmetrically (4).

Position, size and nature of image formed by the spherical mirror

Mirror	**Location of the object**	**Location of the image**	**Magnification, size of the image**	**Nature**	
				Real virtual	**Erect inverted**
(a) Concave	At infinity, i.e., $u = \infty$	At focus i.e., $v = f$	$m << 1$, diminished	Real	inverted
	Away from centre of curvature ($u > 2f$)	Between f and 2f, i.e., $f < v < 2f$	$m < 1$, diminished	Real	Inverted
	At centre of curvature $u = 2f$	At centre of curvature, i.e., $v = 2f$	$m = 1$, same size as that of the object	Real	inverted
	Between centre of curvature and focus : $F < u < 2f$	Away from the centre of curvature $v > 2f$	$m > 1$, magnified	Real	inverted
	At focus, i.e., $u = f$	At infinity, i.e., $v = \infty$	$m = \infty$, magnified	Real	inverted
	Between pole and focus $u < f$	$v > u$	$m > 1$, magnified	Virtual	erect
(b) Convex	At infinity,	At focus,	$m < 1$,	Virtual	erect

	i.e., u = ∞	i.e., v = f	diminished		
	Anywhere between infinity and pole	Between pole and focus	m < 1, diminished	Virtual	erect

1.14. SUMMARY:

1. Lens formula for a thin lens is given by

$$\frac{1}{f} = \frac{1}{v} - \frac{1}{u}$$

Where, u is the object distance, v is the image distance and f is the focal length of the lens.

2. **Lens Maker's formula for a thin lens:** If a thin lens made up of material of refractive index μ_2 is kept in the medium of refractive index μ_1, then its focal length in the given medium is given by

$$\frac{1}{f} = \left(\frac{\mu_2}{\mu_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

Or $$\frac{1}{v} - \frac{1}{u} = \left(\frac{\mu_2}{\mu_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

Where, R_1 is the radius of curvature of first surface of the lens, R_2 is the radius of curvature of second surface of the lens and f is focal length of the lens. If the lens made up f material of R.I. μ is kept in air, then the focal length of the lens in air medium is given by

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

3. **Equivalent focal length:** Equivalent focal length of two thin lenses placed co-axially and separated by a distance x from each other is given by

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} - \frac{x}{f_1 f_2}$$

Or $\quad f = \frac{f_1 f_2}{f_1 + f_2 - x} = \frac{f_1 f_2}{\Delta}$

where $\Delta = f_1 + f_2 - x$ is called the optical separation or optical interval between two thin lenses, f_1 is focal length of first lens and f_2 is focal length of second lens.

The equivalent focal length of two thin lenses in contact is given by

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$$

4. Cardinal points: There are six points in all, called cardinal points in a co-axial system of lenses (or optical system) with the help of which, the position of an object and its image can be obtained.

The following are the cardinal points of an optical system:

(i) Two focal points (F_1, F_2)

(ii) Two principal points (P_1, P_2)

(iii) Two nodal points (N_1, N_2)

Focal points: The focal points of an optical system are pair of points on the principal axis of the system and conjugate to points at infinity.

Nodal points: Nodal points are defined as a pair of conjugate points on the principal axis having unit positive angular magnification.

Principal points: There are two principal planes and two principal points.

The first principal plane in the object space is the locus of the points of intersection of the emergent rays in the image space parallel to the principal axis and their conjugate incident rays in the object space.

Similarly, the second principal plane in the image space is the locus of the pints of intersection of the incident rays in the object space parallel to the principal axis and their conjugate emergent rays in the image space. The point of intersection of first principal plane with the principal axis is called first principal point and the point of intersection of second principal plane with the principal axis is called second principal point. These points ae denoted by P_1 and P_2 respectively.

The distance of the first principal point P1 from the first lens (measured to the right of first lens) is given by

$\alpha = \frac{xf}{f_2}$

The distance of the second principal point P_2 from the second lens (measured to the left of second lens) is given by

$\beta = -\frac{xf}{f_1}$

5. **The deviation produced by a thin lens** is given by

$\delta = \frac{h}{f}$

where h is the height at which the ray strikes from the principal axis and f is focal length of the thin lens.

6. The power of lens is also called the optical power of the lens.

The power of lens is the reciprocal of the focal length of the lens.

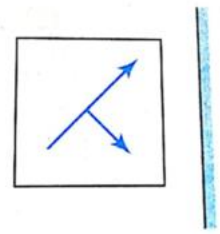
$$P = \frac{1}{f}$$

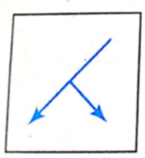
7. **Liner magnification:** It is defined as the ratio of the size of the image to the size of the object.

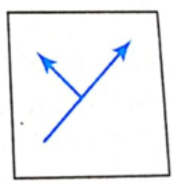
1.15 TERMINAL QUESTIONS:

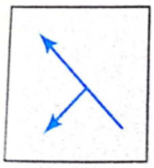
1. Define the terms: optical centre, principal axis and radius of curvature.
2. Define the term lens. Draw the sketch of converging lens.

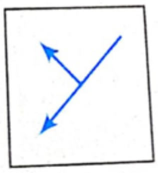
3. What happens when two lenses, one convex and other concave type of same material and having same radii of curvature are joined together?
4. Derive lens formula.
5. Explain what do you mean by an equivalent lens.
6. Define cardinal points of system of co-axial lenses.
7. What do you mean by cardinal points?
8. What do you mean by aplanatic points?
9. What are aplanatic points? Find the aplanatic foci for a spherical refracting surface. Show how they are utilized for the construction of oil immersion objective lenses.
10. What is an eyepiece? Describe in general.
11. Compare the performances of Huygen's and Ramsden's eyepieces.
12. State merits and demerits of Ramsden's and Huygen's eye pieces.
13. What is the range of vision of normal eye?
14. Choose the correct mirror image of object as shown in the following figure.



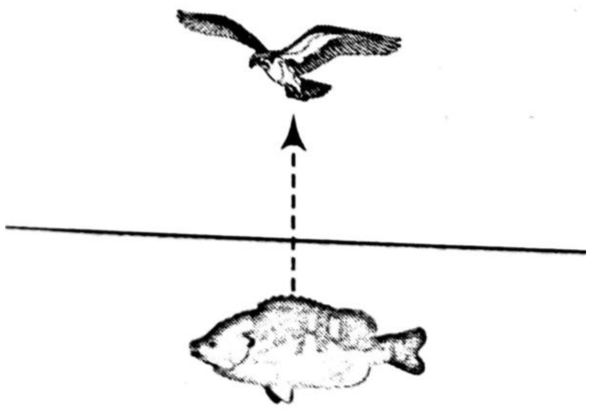
(a) 

(b) 

(c) 

(d) 

15. A fish is vertically below a flying bird moving vertically down toward water surface. The bird will appear to the fish to be



(a) moving faster than its speed and also away from the real distance

(b) moving faster than its real speed and nearer than its real distance

(c) moving slower than its real speed and also nearer than its real distance

(d) moving slower than its real speed and away from the real distance.

16. Mark the correct statement(s) w.r.t. a concave spherical mirror.

(a) For real extended object, it can form a diminished virtual image

(b) For real extended object, it can form a diminished virtual image

(c) For virtual extended object, it can form a diminished real image

(d) For virtual extended object, it can form a magnified real image.

17. Mark the correct statement(s) from the following:

(a) Image formed by a convex mirror can be real

(b) Image formed by a convex mirror can be virtual

(c) Image formed by a convex mirror can be magnified

(d) Image formed by a convex a convex mirror can be inverted

18. A real object is moving toward a fixed spherical mirror. The image

(a) must move away from the mirror

(b) may move away from the mirror

(c) may move toward the mirror if the mirror is concave

(d) must move toward the mirror if the plane mirror is convex

19. A concave mirror forms an image of the sun at a distance of 12 cm from it.

(a) The radius of curvature of this mirror is 6 cm.

(b) To use it as a shaving mirror, it can be held at a distance of 8-10 can from the face.

(c) If an object is kept at a distance of 24 cm from it, the image formed will be of the same size as the object.

(d) All the above alternatives are correct.

20. In column I, types of the mirrors are given and in column II, types of the image for virtual objects are listed, match the following two columns.

Column I		**Column II**	
i.	Concave mirror	a.	only real image
ii.	Convex mirror	b.	only virtual image

iii.	Plane mirror	c.	may be real or virtual image

(a) i. → a.; ii. → c.; iii. → a.

(b) i. → b.; ii. → c.; iii. → a.

(c) i. → c.; ii. → b.; iii. → a.

(d) i. → a.; ii. → c.; iii. → b.

1.16 ANSWER AND SOLUTIONS OF TERMINAL QUESTION

1. Section 1.4
2. Section 1.4, 1.4.1
3. Section 1.4
4. Section 1.5.1
5. Section 1.5.3
6. Section 1.8
7. Section 1.8
8. Section 1.11
9. Section 1.11
10. Section 1.7
11. Section 1.7.1
12. Section 1.7.1
13. Range of vision of a normal human eye is the range of distance for which human eye can see an object clearly. It ranges from infinity to 25 cm i.e., the range between far and near point.
14. (c)
15. (a)

16. (b) & (c)

17. (a), (b), (c), (d)

18. (b), (c), (d)

19. (b), (c)

20. (a)

1.17 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.
2. Optics – Ajoy Ghatak
3. Optics – Principles and Application, K. K. Sharma, Academic Press.
4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.
5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT: 02 LASER AND HOLOGRAPHY

STRUCTURE

2.1 INTRODUCTION

LASER is an abbreviation of "Light amplification by stimulated emission of radiation". It is a device for producing very intense, unidirectional, monochromatic and coherent visible light beams.

2.2 OBJECTIVES

After studying this unit, students were be able to:

- Define, what is LASER & why it is being used?

- Define the terms spatial and temporal coherence.
- Explain spontaneous and stimulated emission.
- Einstein coefficients and its derivation.
- Basic idea of LASER and its component.
- Pumping and Population inversion
- Comparison of LASER light and ordinary light
- Types of LASER
- Applications of LASER
- Holography and its comparison with photography
- Applications of Holography

2.2 TEMPORAL AND SPATIAL COHERENCE

If a wave appears to be a pure sine wave for an infinitely large period of time (or in an infinitely extended space), then it is said to be a perfectly coherent wave.

There are two types of coherence.

Temporal coherence

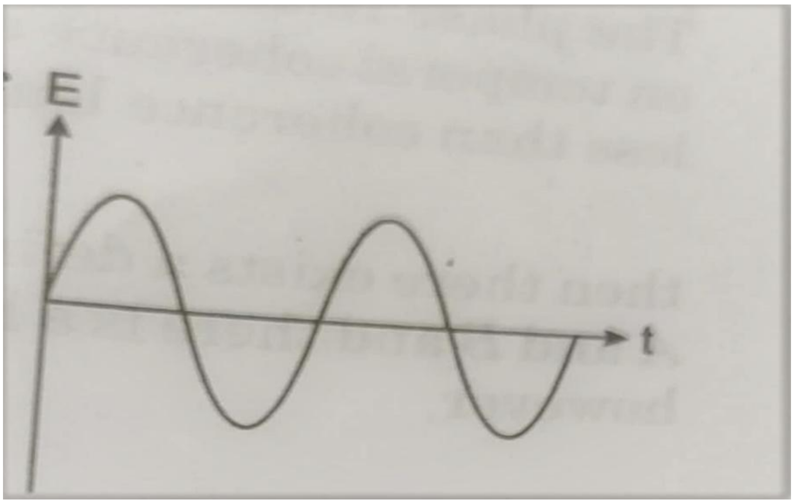


Figure 1:

It is criteria of coherence that depends on upon time. If we plot the light wave, it would appear as shown in fig.1. It is an ideal sinusoidal function of time with constant amplitude at any point while its phase varies linearly with time. But no actual light source limits a perfectly sinusoidal wave. Actually when a light pulse of short duration (of the order of 10^{-10} second for sodium atom) is emitted when an excited atom returns to the initial state. Thus field remains sinusoidal only for 10^{-10}sec. After it phase changes abruptly, as shown in fig.2

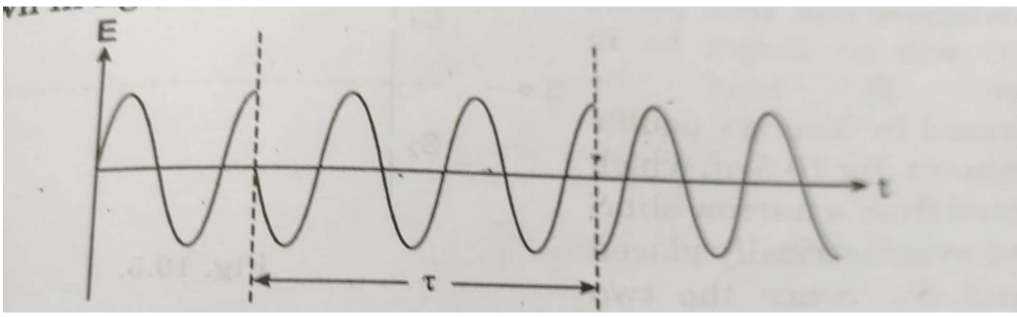


Figure: 2.

The average time interval for which the field sinusoidal is called coherence time or temporal coherence of light beam and is denoted by Ʈ. In this time, the definite phase

Relationship exists. The distance for which the field is sinusoidal is called coherence length and is denoted by L

$$L = ƮC$$

For sodium light L=3 cm, $C=3x10^8$ m/sec

$T=L/C =3cm/3X10^8 cm/sec =10^{-10}$ sec

Spatial coherence

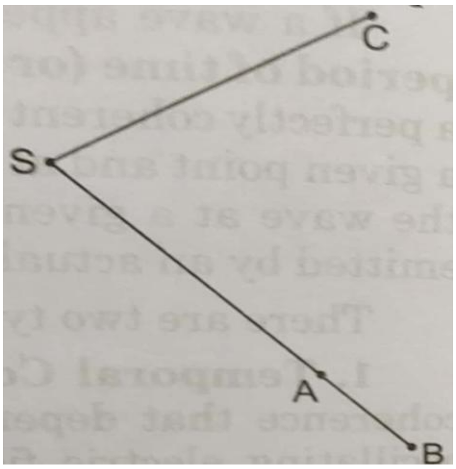


Figure. : 3

If there is a definite phase relationship between the radiation fields at different points in space then there will be high coherence between the points, which is called spatial coherence.

Let S be a point source and two points A and B on a line Joining them with S.(fig.3) The phase relationship between these points depends on temporal coherence and on the distance AB, If AB is less then coherence length L i.e

$$AB<<L$$

Then there exists a definite phase relationship between A and B and therefore is a high coherence between them. If

AB>>L, then there is no coherence between A and B.

If we consider two equidistant points A and C from S,then the waves will reach A and C in exactly the same phase, if S is a true point source. Thus the points A and C have perfect spatial coherence.

2.4 SPONTANEOUS AND STIMULATED EMISSION

Basically, three transition processes can take place when a photon is incident on a system.

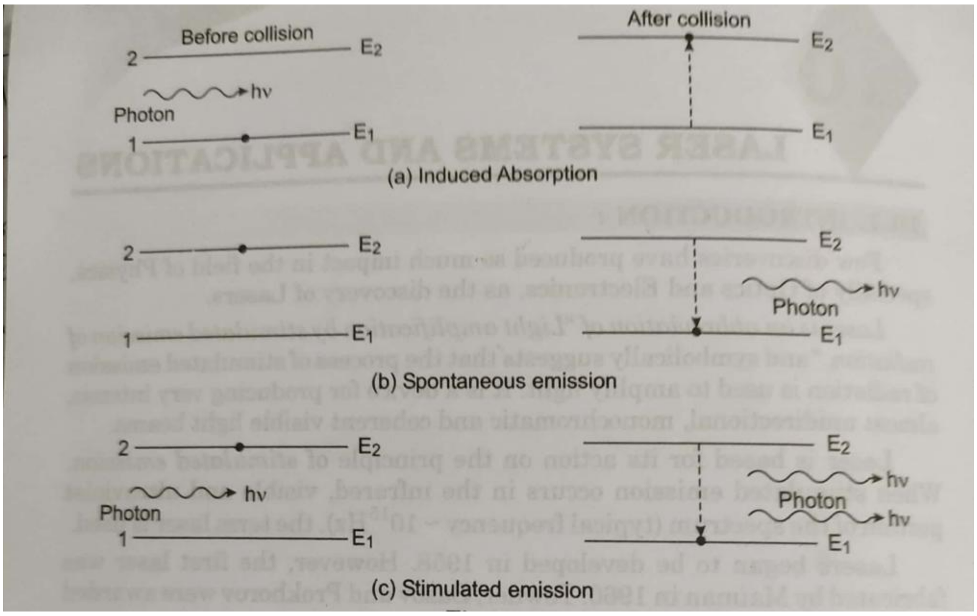


Figure : 4

Absorption, spontaneous emission and stimulated emission.

Consider a simple two level system consists of energy level E_1 and E_2 in an atom, such that $E_2>E_1$ as shown in figure 4. Such an atom can emit or absorb a photon of frequency given by

$$h\nu=E_2-E_1$$

At ordinary temperature, most of the atoms are in the ground state E_1. If a photon of frequency ν is in the ground state E_1 and will be absorption as shown in fig. 4(a)

The probable rate of absorption of radiation of frequency ν effecting tramsition from state 1to state 2 is proportional to the energy density of radiation u(r)

$$P_{12}=B_{12}\, u(\nu) \qquad \text{.................. (1)}$$

Where proportionality constant B_{12} is called Einstein's coefficient of absorption of radiation. Once the atom is in the excited state, it can decay (or drop back to a lower energy state) after a short time typically 10^{-9} to 10^{-3} sec. by two different processes.

2.4.1 Spontaneous Emission

When an atom in an excited state E_2 falls to the ground state E_1, by emitting a photon of frequency $\nu = (E_2-E_1)/h$, the process is known as spontaneous emission and has been shown in fig. 4(b)

In spontaneous emission.

(1) The emitted photon has energy h ν and can move in any random direction.

(2) The photons emitted from various atoms in the assembly have no phase relationship between them.

(3) The rate at which electrons fall from the excited level E_2 to lower level E_1 at energy instant is proportional to the number of electrons remaining in E_2

(4) The transition probability depends only on two energy states.

So radiation given out in spontaneous emission are incoherent.

The probability of spontaneous emission 2→1 is determined purely by the properties of states 2 and 1. Einstein denoted this probability per unit time by

$$(P_{21})_{\text{spontaneous}} = A_{21} \quad \text{----------------(2)}$$

A_{21} is known as Einstein's coefficient of spontaneous emission of radiation.

2.4.2 Stimulated or Induced Emission

When a photon of frequency exactly equal to $\nu = (E_2 - E_1)/h$ is incident on the atom in excited state E_2, then it induces (or stimulates) the atom to move to ground state E_1 by emitting a photon of the same frequency ν. This process is known a stimulated or induced emission as shown in fig. 4(c)

In stimulated (or induced) emission.

(1) For every incident photon, we have two outgoing photons, going in the same direction.

(2) The emitted photons travels in the direction of the incident photon.

Thus the emitted photons have the same frequency of and are in phase with the incident photon.

In this way, we can achieve an amplified as well as an unidirectional coherent beam.

The rate of stimulated emission is proportional to:

(1) The instantaneous number of atoms in the excited state E_2.

(2) The energy density of the incident radiation.

So, the probability of stimulated emission transition2→1 can be written as

$$(P_{21})_{\text{stimulated}} = (B_{21})\, u(\nu) \qquad \text{-(3)}$$

Where B_{21} is called the 'Einstein coefficient of stimulated emission of radiation'.

Thus the total probability of emission transistion from state 2 to drop to state 1 (2→1) is the sum of spontaneous and stimulated emission probabilities given by equations (2) and (3) respectively,

Then $P_{21}=A_{21}+B_{21}\, u(\nu)$ - (4)

2.4.3 Einstein Coefficient and Derivation of Relation Between Them

Einstein's coefficient are

B_{12}- Einstein's coefficient of absorption of radiation

A_{21}-Einstein's coefficient of spontaneous emission of radiation.

B_{21}-Einstein's coefficient of stimulated emission of radiation.

Consider an assembly of atom in thermal equilibrium at temperature T with radiation of frequency ν. The rate of absorption of radiation affecting transition from state 1→2 (Fig 4(a)) is proportional to the energy density of radiation u(ν).

Thus, number of transitions per unit time per unit volume from 1→2 is

$N_1\, P_{12}=N_1\, B_{12}\, u(\nu)$ - (5)

Where N_1 is number of atoms per unit volume in state 1.

The number of spontaneous emission per unit time per unit volume will be proportional to N_2 which represents the number of atoms per unit volume of state 2. So it is equal to $N_2\, A_{21}$.

The number of stimulated emission per unit time per unit volume would depend on N_2 and energy density u(ν). So it is equal to $N_2\, B_{21}\, u(\nu)$

Thus the number of atoms in state 2 that drop to 1 either spontaneously or under stimulation, emitting a photon per unit time per unit volume is

$N_2\ P_{21}=N_2[A_{21}+B_{21}\ u(\nu)\ ]$ -

(6)

In thermal equilibrium, the absorption and emission rates must be equal

$N_1\ P_{12}= N_2\ P_{21}$

$N_1\ B_{12}\ u(\nu) = N_2\ [A_{21}+B_{21}\ u(\nu)]$

$u(\nu)\ [N_1B_{12}-N_2\ B_{21}] = N_2\ A_{21}$

$u(\nu) = N_2\ A_{21}/\ N_1\ B_{12}-\ N_2\ B_{21}$

$u(\nu) = A_{21}/N_1/N_2\ B_{12}-B_{21}$

$U(r) = A_{21}/\ B_{21}\ \ 1/N_1/N_2(B_{12}-B_{21})-1$ - (7)

The equilibrium distribution of atoms among different energy states is given by Boltzman's law, According to which the number of atoms N1 and N2 in energy states E2 and E1 at absolute temperature T are given by

$N_2 = N_0\ e\ \ -E_2/kT$

$N_1 = N_0\ e\ \ -E_1/kT$

$N_2/N_1 =e- (E_2-E_1)/kT$

But $E_2- E_1 = hr$ (Energy of photon emitted or absorbed)

$N_2/N_1 =e- hr/kT$

$N_1/N_2 = e\ hr/kT$

Substing value of N1/N2 in equation 7 we get

$U(r) = A_{21}/B_{21}\ \ 1/e\ hr/kT\ (B_{12}/B_{21})-1$

Comparising this with planck radiation formula

$U(r) = 8\pi\ hr3/C3*1/e\ hr/kT-1$

We get

A_{21}/B_{21} = 8π hr3/C3 - (8)

And B12/B21 =1 - (9)

From 9 we found that,

$B_{12} = B_{21}$

i.e the probability of stimulated emission is equal to the probability of induced absorption

from 8 we observe that

$A_{21}/B_{21} \partial r_3$

i.e. the ration of spontaneous emission and stimulated emission is proportional to r3. It means that the probability of spontaneous emission increases rapidly with the energy difference between the two states.

2.5 BASIC IDEA OF LASER AND ITS COMPONENTS

A LASER is a device to a produce an intense, highly concentrated, monochromatic, unidirectional and highly coherent beam of light. Typical operating frequency of LASER is 10^{15} HZ in visible region.

An excuted atom emits energy in the form of light when the secuted electrons in the atom drop back to a lower state. Such an emission of light is known as spontaneous emission in which the radiation is emitted in all directions and has energy equal to energy difference between the excuted state and the lower energy states. If however during the excuted state of the atom,

It is seposed to a matching photon (which has exactly the same frequency as that emitted by this atom in spontaneous emission), it stimulates the atom in secited state to decay by emitting a photon equal in frequency to the ine which stimulated it. Such a photon travels in the direction of incident photon and for energy incident photon, we have two outgoing photons.

Thus light amplification takes place. In order to have a unidirectional, coherent beam, we should have more stimulated emussion than spontaneous emission. Thus is achieved by population inversion, so that the population of atoms in higher energy state E2 is more than population of atom in lower energy state E1 i.e.

$E1<E2$ but $h2>h1$

The procedure adopted to achieve population inversion is called pumping, in which the atoms from secited state first decay spontaneously to a metastable state where they struck by matching photons resulting in a chain of stimulated emission and consequently giving rise to a coherent, highly intense beam of photons travelling in the direction of incident beam.

2.5.1 Pumping

The process of achieving population inversion is called pumping a LASER and producing population inversion. Some of the commonly used methods for pumping is

(1) Optical pumping

(2) Electric discharge

(3) Inelastic atom-atom collision

(4) Direct conversion

(5) Chemical reactions

1. Optical Pumping :

If luminious energy is supplied to medium for causing population inversion, then pumping is called the optical pumping. In optical pumping the luminious energy usually comes from a light source in the form of short flashes of light.

This method was first used in Ruby LASER by Maeman and at present being used in solid state Laser.

2. Electric Discharge :

The pumping by electric discharge is preferred in gaseous- ion Laser (e.g. Argon-ion Laser). In discharge tube when a potential difference (p.d.) is applied between cathode and anode; the electrons collide with atoms of the active medium, ionize the medium and raise it o the higher level.This produce the required population inversion. This is also called direct- electron excitation.

3. Inelastic Atom- Atom Collisions :

In electric discharge on type of atoms are raised to their excited states. These atoms collide inelastically with another type of atoms, It is these latter atoms which provide the population inversion needed for laser emission, The sample is helium neon laser.

4. Direct Conversion

A direct conversion of electrical energy into radiant energy ommurs in light emitting diodes (LED's) . The sample of population inversion by direct conversion occurs in semiconductors lasers.

5. Chemical Conversion:

In a chemical laser energy comes from a chemical reaction without any head for other energy sources, For example hydrogen can combine with fluorine to form hydrogen- fluoride

$H_2+F_2 \rightarrow 2HF$

This reaction is used to pump a CO_2 laser achieve population inversion.

6. Mechanism of Optical Pumping:

In optical pumping, the population inversion is brought about by three level scheme. Atoms excited from level E_1 level is short lived state. These short lived atoms do not fall directly in level E_3, because direct transition in level E_3 to level E_1 is forbidden by selection rules, but transition from E_3 to E_2 (meta stable state) is allowed having more lifetime than E_3, when the atoms of the substance are irradiated with an excitation frequency $r=E_3-E_1/h$, the atoms are excited to the state E_3 by the process of stimulated absorption. Some of the atoms decay spontaneously to metastable state E_2, where they live for much longer time of 10_3 second as compared to 10_8 second for the short lived state E_3. This enventually leads to situation that state E_2 has more atoms than states E_1, which means populations inversion has been achieved.

2.5.2 Population Inversion

Suppose we have large number of atoms, say no in thermal equilibrium, then their distribution in different energy states obeys Maxwell-Boltzmann statistics.

If we assume that at temperature T K, the instantaneous populations in energy state E_1 and E_2 are n_1 and n_2 respectively, then

$N_1= n_0$ e $–E_1/kT$

$N_2= n_0$ e $–E_2/kT$

Then $n_2/n_1=$ e $–(E_2-E_1)/kT$

Where $k = 1.38 \times 10^{-23}JK^{-1}$ Joule, K is Boltzmann constant:

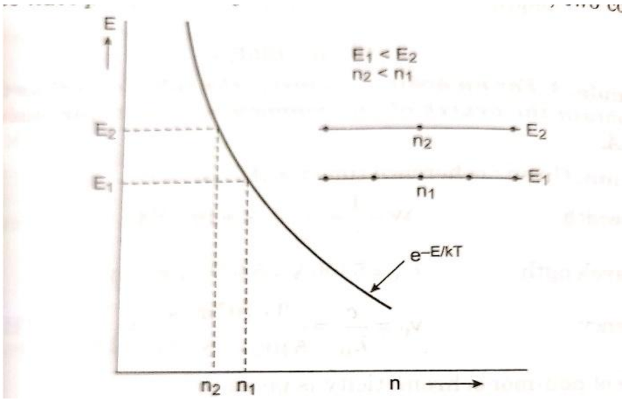


Figure : 5

As $E_2 > E_1$, So $n_2 < m_1$ this means the number of atoms in higher energy state is less than number of atoms in higher of atoms in lower energy states. Under such circumstance, the probability of stimulated emission is much less than the probability of stimulated emission is much less than the probability of spontaneous emission. But for producing coherent and amplified Laser light probability of stimulated emission must increase

Thus basic requirements of a laser is to have predominantly stimulated emission. For it two conditions must be satisfied.

1. The higher energy state should have a longer mean life time i.e. it should be a metastable state.
2. The number of atoms in the higher energy state E_2 must be greater than that in E_1 i.e.

$$E_1 < E_2 \text{ but } n_2 > n_1$$

As shown in figure 6

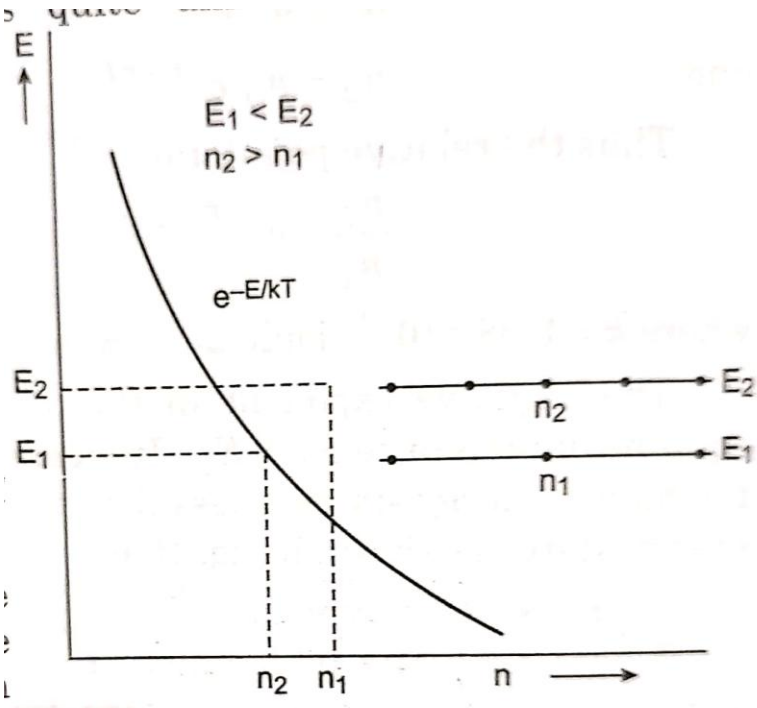


Figure : 6

This condition of $n_2>n_1$ is quite unnatural, because for any equilibrium state n_2/n_1 is less than unity. If by some means, a large number of atoms made available in the higher energy state is greater than that in lower energy state is called population inversion when this is achieved, the emitted photons having same frequency and phase interact with other atoms and thus a sort of chain emission states in which all the photons have same frequency and phase and travel in the same direction.

Thus the number of identical photons goes on multiplying by repeated stimulated emission and we get a highly intense, monochromatic, coherent and unidirectional beam from such a source as shown in figure 7

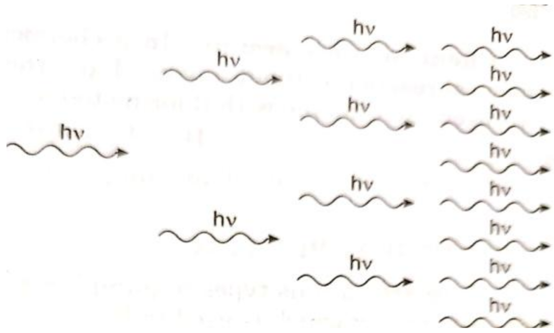


Figure : 7

2.6 TYPES OF LASER

LASER system may be classified in two ways.

1. On the basis of output beam

On this basis Lasers are of two types.

(1) Continuous wave Lasers: The Lasers which give output in the form of continuous wave(or CW) Lasers. The examples are Nd. YAG Laser, He- Ne Laser, Argon ion Laser Co2 Laser etc.

(2) Pulsed Laser: The Lasers which give output in the form of pulses are called pulsed Lasers. The example of pulsed lasers are Ruby Laser, Nd-Glass Laser, Nitrogen Laser etc.

2. On the basis of active medium

According to this Lasers are three types

(1) Solid state Lasers : If the active material is in the form of solid state, the Laser is said to be solid state Laser. The example are Ruby Laser, Nd- YAG Laser, Nd-Glass Laser Neodymium Lasers etc.

(2) Dye(or liquid) Lasers: If the active material is in the form of liquid, the Laser is said to be liquid (or dye) Laseractive medium is formed by solution of certain organic dyes dissolved in liquid such as alcohol and water .

These dyes belong to following Lasers

- Polymethine dye
- Xanthene dye
- Courmarine dyes
- Scintillation dyes

3. **Gas Lasers :** If the active material is the form of gas or vapour, the Laser is said to be a gas Laser. These are three types of gas Laser depending on the nature of active medium. They are

   a) Atomic Lasers: Examplws is He-Ne Laser.

   b) Ionic Laser: The examples are Argon-ion and Krypton-ion Laser.

   c) Molecules Laser: The example are carbon- dioxide Laser, Eximer Lasers, Nitrogen Laser etc.

2.6.1 Ruby of Laser

It is the first working laser developed by Maximum in 1960. It is a solid state Laser and makes use of the three level scheme of population inversion.

It consists of three essential parts.

- The working Material: The working material is a ruby crystal which belongs to a family of gems. Actually it is aluminium oxide (Al_2O_3) doped with 0.05% chromium oxide (Cr_2O_3). Thus some of the aluminium atoms in the crystal lattine are replaced by cr++ ions which give ruby a pink colour.
- The pesonant cavity: It is a speciallu prepaced ruby cylinder in which light intensity can be build up by multiple reflections. For this purpose a

ruby rod 20 to 30 cm. long and 0.5 to 2 cm in diameter is grown. The crystal is cut and polished,So that the ends are flat and parallel with the end planes perpendicular to the axis of the rod. One of its ends is completely silvered to become fully reflecting while the other is only partially silvered so that an intense beam can emerge out of it.

4. The optical pumping system

   The optical pumping is done by an external source. In general, a helical xenon flash tube acts as a pump. It is wound round the glass envelope surrounding the ruby rod through which liquid nitrogen is circulated to keep the rod cool.

   All these parts of a ruby laser is shown in figure 9

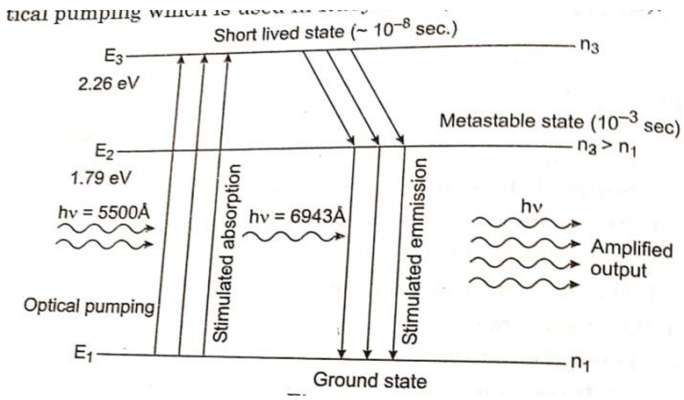


Figure : 9

Operation : The active material in ruby laser is cr++ ions. The energy level diagram for these ions is shown in figure 8.

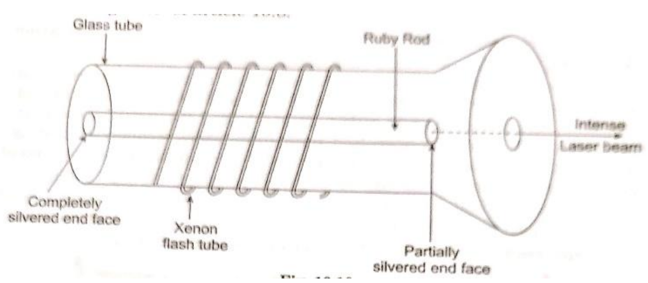


Figure : 8

The optical pumping results, when a flash of light from Xe flash tube falls upon the ruby rod. As a result a green yellow radiations of = 5500A^{o} are absorbed by the chromium ions, which are exicted from state E1 to energy state E_3. The cr++ ions exicited to shorter lived 2.26 ev, E_3 state , jump to 1.79 ev metastable state E_2 by losing 0.47ev of energy in the first step. These radiations are non radiative transitions the excess energy is absorbed by the lattine and does not appear in the form of electromagnetic radiations.

Since metastable state has very long life time(~ 3×10^{-3} sec) as compared to 10^{-8} sec of E_3, The number of states in E_2 are increasing and ultimately exceeds than ground state E_1 and hence a population inversion is established between metastable state E_2 and ground state E1.Such a system with population inversion is very instable. When an atom decays spontaneously from the metastable state E_2 to ground state, photon of energy 1.79 ev corresponding to red light of λ=6943 A^{o} is produced. This photon travels through the ruby rod and if it moves parallel to the axis of the crystal, is reflected back and forth between the two ends of the crystals, one of which is fully and other is partially silvered to acts as resonator. It stimulates an excited atom to emit a fresh photon inphase with the stimulating photon. The process is repeated again and again, because the photons undergo multiple reflection from the silvered ends. Thus a chain reaction or avalanche effect is produced till the beam becomes sufficiently intense to emerge out of the partially silvered end of the crystal. Thus highly intense, coherent, monochromatic and unidirectional beam is obtained.

2.6.2 Helium - Neon Laser

A Continuous Wave Laser

Two important advantages of a ruby Laser are

i. Output beam is not continuous.

ii. Ir requires a large amount of optical energy to create population inversion.

Helium-Neon Laser is a gas laser which emits light continuously, rather then in pulses. It uses in mixture of helium and neon gases. It differs in operation from ruby laser, because four energy levels are involved in it- one in helium and three in neon. The excitation of helium and neon atoms is obtained by means of high frequency electromagnetic field. The energy is transferred to the atoms of the gas by electron impact and collisions between atoms.

Construction: A typical He-Ne laser is shown in fig.10

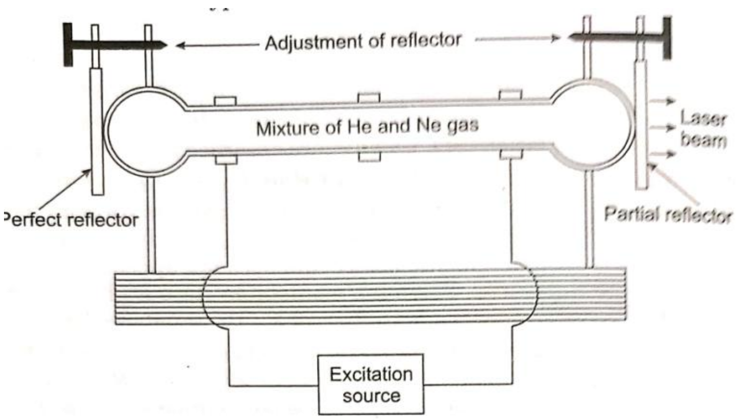


Figure : 10

It consists of

i. A nearly 1 metre narrow quatz tube containing He and Ne in the ratio 7:1 at a pressure of 1 torr (1mm of Hg).

ii. An excitation source for creating a discharge frequency potential difference, such as that obtained three metal bands around the outside of the tube can be adjusted to a high degree of parallelism. At on end of the tube, however, the mirror is fully silvered (perfect reflector) while at the other end is partially silvered (partially reflector).

Operation

When the electromagnetic energy is injected into the tube through metal bands, helium atom, are excited to metastable states by means of collision with the electrons in the tube. The excited helium atoms collide with the unexcited neon atoms, thereby transplanting inversion in Ne atoms. Since ‘He’ has metastable levels at almost the same energies a resonant exchange of energy and a resultant increase in population of the corresponding Ne levels. The laser action takes place in neon atoms, helium in the mixture serves the only purpose to enhance the excitation process.

The Ne levels important for 2 s laser action are shown in fig 11.

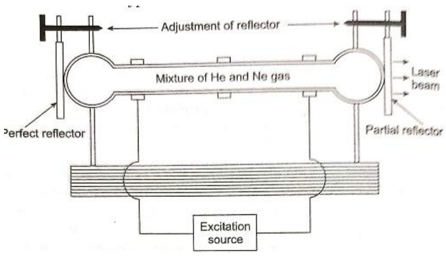


Figure : 11

The strongest emission line occurs between a 2 s and a 2p level with a wavelength 11523 A^{o} which lies in the infrared.

Another important transition takes place between a 3s level and 2p level giving off photons of wavelength 6328A^{o} which lies in the red part of the visible spectrum. A third transition between 3s and 3p levels gives photon of wavelength 3.39Hm.

Light waves emitted parallel to the tube axis bounce back and forth between silvered ends and stimulates emission of the same frequency from other excited neon atoms.

Thus the photons get multiplied and the initial beam is built up into a powerful, coherent, parallel beam and emerges through the partially silvered end of the tube.

2.6.3 Semiconductor Lasers

Semiconductor are materials which have electrical conductivities between those of conductors and insulators. A semiconductor laser developed in 1962 , these devices differ from the solid, liquid and gas lasers in several aspect. They have remarkably small size, rehibit high efficiency and opsate at lower power. Thus they provide a portable and easily controlled source of lower power coherent radiation, particularly suited for fibre optics communication systems.

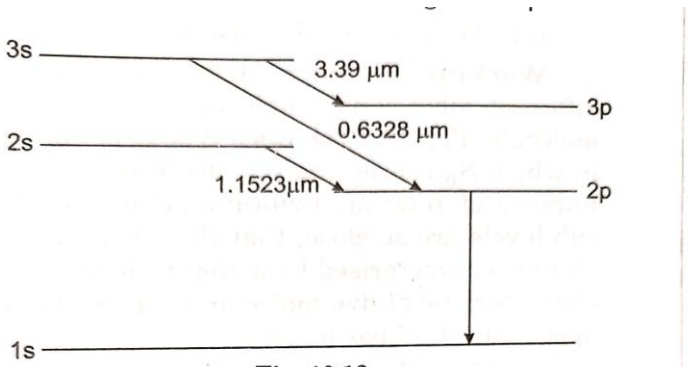


Figure : 12

Population invesion in semiconductor is achieved by using a P-N Junction diode, heavily doped with donars and acceptors. When no barriers exists across the depletion, a potential Junction diode and no current flows. When P-N is connected to positive terminal of battery and N type semiconductor is connected to negative terminal of battery, then holes from P region is injected into N region and electrons from N region is injected into P region , so electrons minority charge carriers and combine with there counter parts. On increasing the bias, current increases, when threshold current is reached, a population inversion is created close to the Junction between the filled level near the bottom of conductor band and empty levels near the top of valence band . Hence light amplification is obtained in thus region leading to a monochromatic and highly directional beam of light from the Junction. Fig.13 shows a semiconductor laser using Ga As in the form of diffused P-N homo Junction

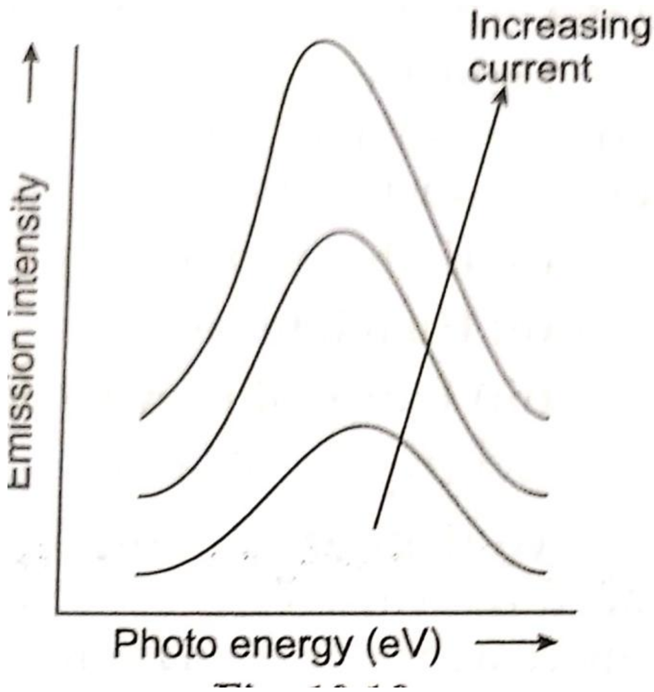


Figure : 13

The surfaces are coated to increase reflectively. Remaining sides of diode are made rough to avoid the leakage of laser beam. The threshold current density increases with temperature for a nondestructive P-N Junction, the typical value being 500 A mm^2 at room temperature.

2.7 COMPARISON OF LASER LIGHT AND ORDINARY LIGHT

LASER LIGHT	ORDINARY LIGHT
1. Laser light is	1. It is a mixture of

monochromatic (having single colour). 2. Laser light shows directional and highly consistent distribution. 3. It has a focused beam in which all photons move at the same wavelength and same direction. 4. Laser have narrow frequency range and directional distribution. 5. Laser light are spectrally purse. 6. In constrast, Laser light is emitted from a narrow beam, since energy is concentrated in narrow area, looking at the Laser light with naked eye can damage the eye.	electromagnetic waves of different wavelength. 2. It is non- directional and inconsistent, which means it travels without follow any direction. 3. Ordinary light has a wide spectrum of light that moves irregularly at different wave lengths. 4. Example of ordinary lights sunlight fluorescent and incandescent etc. 5. Ordinary light does not contain photons of the same frequency. 6. The intensity of light decreases rapidly as it travels a long distance.

2.8 APPLICATIONS OF LASER

Some of important applications of Lasert are listed below:

i. <u>In surgery</u>: A laser beam is safely employed for welding a detached antenna of the eye by focusing it on the spot to be welled. Intense and powerful laser beam has also been used to remove eye tumour.Intense laser beam also being used in treatment of human and animal cancer Laser acts as a sharp knife and its lqcalized heating seals the blood vessels instantly.

ii. In Radio Communication and space Exploration : Laser light being highly coherent can be modulated to transmit hundreds of message at a time on radio and televersion . It can accommodate, in a single frequency and, greater number of channels of carrier wave with enough bandwith. It is also being used for underwater communication, because it is not easily absorbed by water.

Narrow angular spread of the laser beam is being utilized for communications with earth satellites and rockers to the moon and other planets.By using Laser beam and principle of reflection earth-moon distance can be measured accurately.

1. **In Industry**

   Lasers have got wide industrial and chemical applications. A laser beam when focused, produces extremely high temperature. It can, so being used in melting and vaporizing metals and drilling holes in diamonds and hard steels.

2. **For Military Purposes**

   Laser beam is being used for dffective and automatic control and guidance of rockets and satellites. When used in Radars, it can be employed to destroy a reoplanes and missiles. Focussed Laser beams would become the legendary ' death ray' of science fictions that would beam every thing standing in the way of the beam.

3. **In Basic Scientific Research**

Laser beam provides a new source of exploring the molecular structure and nature of chemical reactions. It is very important in the study of Raman Spectroscopy, perversion measurement of length and accurate determination of velocity of light.

4. **Barcode scanners**

Supermarket scanners typically use helium-neom lasers to scan the universal barcodes to identify products. The laser beam bounces off rotating mirror and scans the code, slanding a modulated beam to a light detector and then to a computer which has the product information stored.

5. **In Laser Printing**

The laser printing has becomes dominant mode of printing in offices. The laser is focused and scanned drum, where it produces a charge pattern, which mirrors the material to be printed. Thus drum then holds the particles of the toner to transfer to paper which is rolled over the drum in the presence of heat.

6. **In CD's and Optical Discs**

The detection of the binary data stored in the form of pits on the compact disc is done with the use of a semiconductor laser.

7. **In Laser cooling**

The use of lasers to achieve extremely low temperature has advanced to the point that temperature of 10 9 K have been reached.

8. **In Holography**

Laser have been used for obtaining very sharp images of moving objects. Thus the use of lasers has made holography (three dimensional photography) possible.

2.9 HOLOGRAPHY AND HOLOGRAM

Let us discuss below :

2.9.1 What is Holography

Images of objects are generally obtained by using photographic method. In this method a lens focuses the light reflected from a three dimensional object on to a photographic film. Where a two dimensional image of the object is foremed. A negative is first obtained by developing the film and then a positive is obtained through printing. The positive print is a two dimensional record of light intensity received from the object. Thus it contains information about the square of amplitude of light wave that produced the image but the information about phase of the is not recorded and lost.

In 1948 Dennis Gabor outline a two step lensless imaging process. It is a new technique of photographing the objects and is known as wave front reconstruction. The technique is also called holography. The word holography is formed by combining parts of two Greek words: holos' meaning 'whole' and 'graphein' meaning " to write". Thus holography means writing the complete image. Holography is actually a recording of interference pattern formed between two beams of coherent light coming from the same source. In this process both the amplitude and phase components of light wave are recorded on a light sensitive medium such as photographic plate. The recording is known as a hologram. Holography technique became a practical proposition after the invention of laser.

2.9.2 Principle of Holography

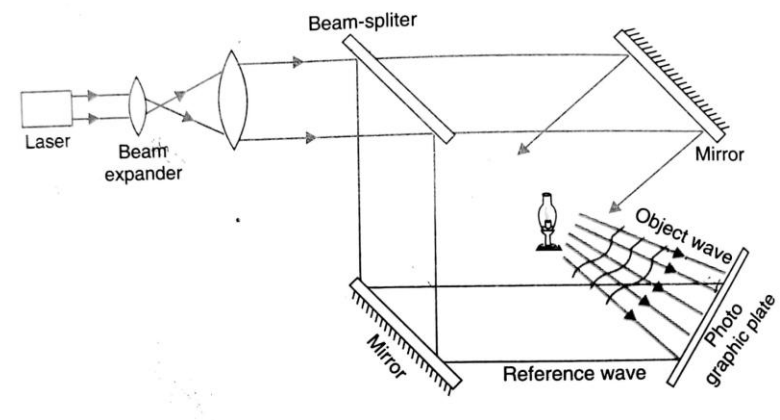
Holography is a two step process. First step is the recording of hologram where the object is transformed into a photographic record and the second step is the reconstruction in which the hologram is transformed into the image.

Unlike the conventional photography, lens is not required in either steps. A hologram is the result of interference occurring between two waves, an object beam which is the light scattered off the object and a coherent back ground, the reference beam, which is the light reaching the photographic plate directly. Leith and Upatmicks used reference beam at an effect angle. That made possible the recording of holograms of there dimensional objects.

The off axis arrangement for generating and viewing holograms is described below.

Recording of the Hologram :

In the off axis arrangement a broad laser beam is divided into two beams, namely a reference beam and an object beam by a bem splitter. Fig.14.



Recording of the hologram

Figure : 14

The reference beam goes directly to the photographic plate. The second beam of light is directed onto the object to be photographed. Each point of the object scatters the incident light and acts as the source of spherical waves. Part of the light, scattered by the object, travels towards the photographic plate. At the photographic plate the innumerable spherical waves from the object combine with the plane light wave from the reference beam. The sets of light waves are coherent because they are from the same laser.

They interfere and form interference fringes on the plane og the photographic plate. These interference fringes are a serviles of zone plate like rings, but these rings are also superimposed, making a complex pattern of lines and swirls (Fig.15).

Figure : 15

The developed negative og these interference does not contain a distinct image of the object but carries a record of both the intensity and the relative phase of the light waves at each point.

Reconstruction of the image :

Whenever required, the object can be viewed . For reconstruction of the image, the hologram is illuminated by a parallel beam of light from the laser (Fig.16.).,

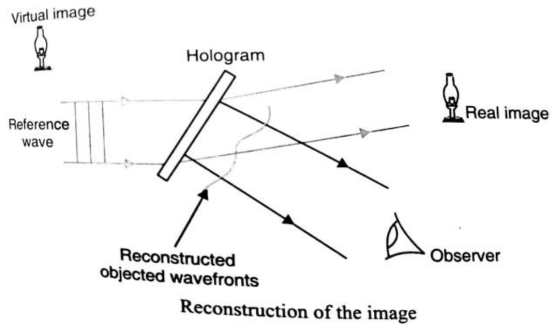


Figure : 16

Most of the light passes straight through, but the complex of fine fringes acts as an elaborate diffraction grating. Light is diffracted at a faisly wide angle. The diffracted rays form two images: a virtual image and a real image occupied by the object and is sometimes called as of the hologram. Since the light rays pass through the point where the real image is it can be photographed. The virtual image is hologram only for viewing. Observer can move to different positions and look around the image to the same extent that he would be able to were he looking directly at the real object. This type of hologram is known as transmission hologram because the image is seen by looking throught it. The three dimensional imagr is seen suspended in mid ray at a point which corresponds to the position of the real object which was photographed.

2.9.3 Hologram and its important properties

Holograms

Holograms are true three dimensional images. This is evidenced by the fact that one can move his head whole viewing the image and it in a different perspective. It reveals part of the image which was hidden at another viewing angle.

For example, three images are shown (Fig.17).

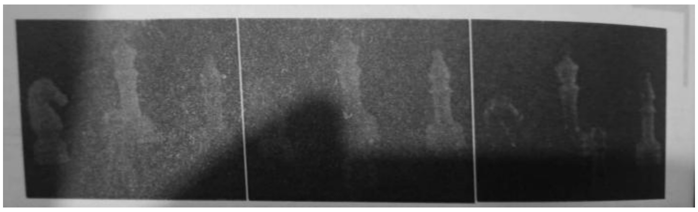


Figure : 17

They are from the same hologram but are obtained by looking through the hologram at different angles. Note that the pawn appears in different perspective is front of the king behind it.

Orthoscopic and Pseudoscopic Images :

A hologram reconstructs two images, one real and other virtual which are exact replicas of the object. However, the two images differ in appearance to the position as the object and has the same appearance of depth and the parallex as the original three dimensional object. The virtual image appears as if the observer is viewing the original object through a window defined by the size of the hologram. This image is known as Orthoscopic image.

The real image is also formed at the same distance from the hologram, but in front of it. In the real image, however the scene depth is inverted. This is due to fact that the corresponding points on the two images (virtual and real) are located at the same distances from the hologram. The real image is known as pseudoscopic image and does not give a pleasing sensation as we do not come across objects with inverted depth in normal life.

Important Properties of Hologram

1. In an ordinary photograph each region contains separate and in dividing part of the original object. Therefore destruction of a portion of a photographic image leads to an irrepairable loss of information corresponding to the destroyed part. On the other hand,

in a hologram each part contains information about the entire object. From a small part of the hologram, the entire image can be reconstructed only with a reduced clarity and definition of the image.

2. It is not useful to record several images on a single photographic film. Such a record can not give information about any of the individual images on the other hand, several images can be recorded on a hologram. Thus the information holding capacity of a hologram is extremly high. For example a 6×9 mm photograph can hold one printed page, where as a hologram of the same size can store upto 300 such pages.
3. On a hologram information is recorded in the form of interference pattern, The type of the pattern obtained depends on the reference beam used to record the hologram. The information can be decoded only by a coherent wave identical to that of the reference wave.
4. The reconstruction of the image of the hologram can be done with reference beam. If the wavelength λ of the reconstructing beam is greater than that λ_0 of the reference image. The magnification will be proportional to the ratio of the two wavelengths.

2.9.4 Comparison of Holography and Photography

Record a hologram is different from taking a photograph.

i. In hologram, it is necessaty that a coherent light source like a laser is used to illuminate the object .
ii. There is a second beam of coherent light, which strikes the film on which the hologram is to be recorded. This is called the reference beam. The reference beam and object beam overlap at the surface of the film and they form an interference pattern.

iii. A high resolution photographic film is used for recording the fine patterns.

iv. Holographic recording is more sensitive to movement and vibrations compared to photographic recording.

The fundamental difference between a hologram and an ordinary photograph is like this.

i. In a photograph the information is stored in an orderly fashion: each point in the object relates to a conjugate point in the image. In a hologram there is no such relationship, light from every object point goes to the entire hologram, This has two main advantages. As the observer moves sideways in viewing the hologram, the image is seen in three dimensions.

ii. If the hologram were shattered or cut into small pieces, each fragment would still reconstruct the whole scene, not just part of the scene.

2.9.6 Applications of Holography

Holography can be used for a broad range of applications in different fields. It is not possible to desirable all of them only some typical applications are discussed here.

1. Security : One of the most popular use of hologram is security and products authentication. For example in identification of documents including credits card, debit card, phone cards, driver licenses etc.
2. Three dimensionsal photography : One of the most obvious applications of holography is the production of a three dimensional photograph, with the distance and orientation of each point of the object recorded in the image.
3. Microscopy : Holography can be used in techniques of microscopy. It is possible to obtain a magnified image of an object by using principle of holography.

4. Character recognition : Holography can also be used for character recongnition. Holography principle can be used to identify finger prints etc.
5. Data Storage : Holograms can be used for data storage devices and hence are of much use in computer technology. A large amount of information 10 12 letters/digits can be stored n a cubic cm of a volume hologram.
6. Photolithography :Holography is used in the production of photographic masks used to produce microelectronic circuits.
7. Holographic projection : This projection is being used to display flight information at the pilot's eye level in an airplane cockpit.
8. Holographic interferometry : One of the most important applications of holography has been in interferomatry. Holographic interferometry is an optical technique to visualize in a dark environment small deformation (200nm to 100km) of objects. It is applied to objects which are placed in a vibration free set up. Holographic interferometry is used in vibrational anqlysis, structural aralysis, strain and strss evaluation. There are three basic methods of holographic interferometry. They are known as real time, multiscope and time average holography.
9. A coustiacal Holography : It is easy to produce coherent sound waves. Therefore holograms can be made using ultrasonic waves initially and then visible light can be used for reconstruction of the visual image. Light waves cannot propogate considerable distances in dense liquids and solids whereas sound waves can propogate through them. Thus a three dimensional acoustical hologram of an opaque object can observed . Such techniques will be highly useful in the fields of medicine and technology.

10. Medical Applications of Holography : Holographic technique is also used in various medical applications, few of them are ophthalmology, Holographic Endoscope, Diffractive bifocal Intraocular Lens, Holography in Orthopedics.

2.10 SUMMARY

A laser is a device that emits a beam of coherent light through an optical amplification process. There are many types of lasers including gas lasers, fiber lasers, solid state lasers, dye lasers, diode lasers and excimer lasers. All of these laser types share a basic set of components.

2.11 TERMINAL QUESTION

1. What is meant by holography ?
2. What is a Hologram ?
3. Write the Principle of holography.
4. Explain different types of LASER.
5. Write down application of LASER.

2.12 ANSWER AND SOLUTION OF TERMINAL QUESTION

1. Section 2.9.1
2. Section 2.9.3
3. Section 2.9.2
4. Section 2.6
5. Section 2.8

2.13 SUGGESTED READING

1. A Textbook of Optics Dr. N. Subrahmanyam Brij Lal & Dr. M. N. Avadhanulu
2. Optics – Ajoy Ghatak
3. Optics – Principles and Application, K. K. Sharma, Academic Press.
4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.
5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT: 03 FIBER OPTICS

STRUCTURE

3.1 INTRODUCTION

In 1870 John Tyndall, a British physicist demonstrated that light can be guided along the curve of a stream of water. Due to total internal reflection light gets confined to the water stream and the stream appears luminous. A luminous water stream is the precursor of an optical fibre. In 1950's the transmission of images through optical fibres was realized in practice. Hopkins and Kapany developed the flexible fiberscope, which was used by medical world in remote illumination and viewing the interion of human body. It was Kapany who coined the term fiber optics. In 1969, it had been established that light could be guided by a glass fibre, but the fibers available at that time have very large attenuation of light passing through them. In 1970 Corning Glass works produced low less glass fibre. The invention of solid-state laser S in 1970 made optical communication practicable. Commercial communication systems based on optical fibers made their appearance by 1977. A part from are widely used in other areas. Fiberscopes made of optical fibers are widely used in a variety of forms in media diagnostics. Sensors for detecting electrical, mechanical thermal energies are made using optical fibres.

Fibre optics s a technology in which signals are converted from electrical into optical signals, transmitted through a thin glass fibre and reconverted into electrical signals.

3.2 OBJECTIVES

After studying this unit, students will be able to define.

- What is Fiber optics & why it is being used?
- Explain constructions and materials used in fiber optics.
- Explain principle of fiber optics.
- Explain the process of propagation of light in optical fiber.
- Explain advantage and disadvantage of fiber optics.
- Explain to various parameters used in fiber optics.
- Explain classification of various fibers.
- Explain attenuations and dispersion in optical fiber parameters.
- Explain the use of couples, splices and connector in fiber optics.

3.3 CONSTRUCTIONS AND MATERIALS USED IN FIBRE OPTICS

Construction and materials. An optical fiber (or fiber in British English) is a flexible, transparent fiber made by drawing glass (silica) or plastic to a diameter slightly thicker than that of a human hair silica (sand siO_2) is a available in abundance in nature.

Optical fibers are used most often as a means to transmit light between the two ends of the fibre and find wide usage in fibre- optic communications, where they permit transmission over long distances and at higher bandwidths (data transfer rates) than electrical cables.

An optical fibre is a cylindrical wave guide made of transparent dielectric (glass or clear plastic) which guides light waves along its length by total internal reflection. It is as thin as human hair, approximately 70 µm or 0.003-inch diameter (Note that a thin strand of a metal is called a wire and a thin strand of dielectric materials is called a fibre.)

3.4 PRINCIPLE OF OPTICAL FIBRE

The propagation of light in an optical fibre from one of its ends to the other end is based on the principle of total internal reflection. When light enters one end of the fibre, it undergoes successive total internal reflections from sidewalls and travels down the length of the fibre along a Zigzag path, as shown in fig 1.

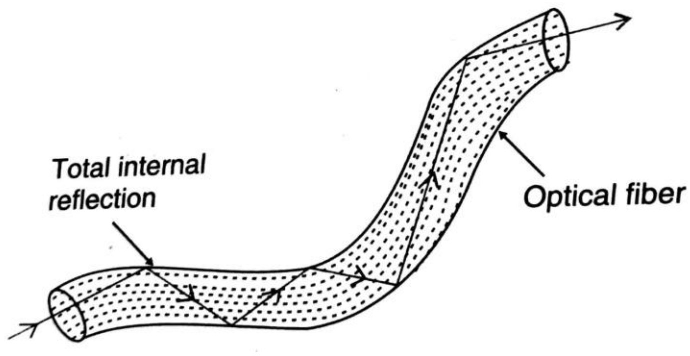


Figure: 1

A small fraction of light may escape through sidewalls but a major fraction emerges out fraction exist end of the fibre as shown in fig.1. Light can travel through fibre even if it is bent.

Fig. 2.side view and cross sectional view of a typical optical fibre.

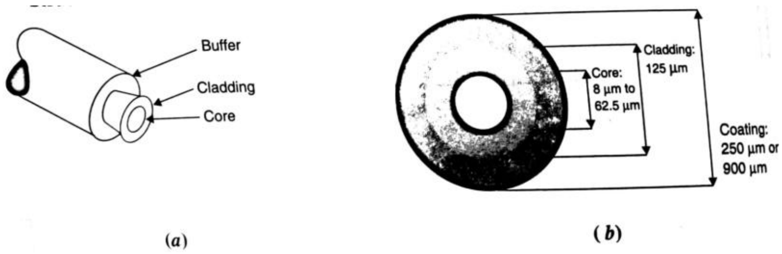


Figure: 2

A practical optical fibre is cylindrical in shape (Fig. 2(a)) and in general has three coaxial regions (Fig 2 (b)).

1. The innermost cylindrical region is the light guiding region known as core. The diameter of the core is in general of the order of 8.5 µm to 62.5 µm.
2. It is surrounded by a casual middle region known as the cladding. The diameter of the cladding is of the order of 125 µm. The refractive index of cladding (n_2) is always lower than that of core (n_1). Light launched into the core and striking the core to cladding interface at an angle greater than critical angle will be reflected back into the core. Since the angle of incidence and reflection are equal, the light will continue to rebound and propagate through the fibre.
3. The outermost region is called the sheath or a protective buffer coating. It is a plastic coating given to the cladding for extra protection.
4. Cladding, reduces the cone of acceptance and increases the rate of transmission of data.

5. A solid cladding, instead of air, also makes it easier to add other protective layers over the fibre.

Optical Fibre System

An optical fibre is used to transmit light signals over long distances. It is essentially a light transmitting medium, its role being very much similar to coaxial cable or wave guide used in microwave communications optical fibre requires a light source for launching light into the fibre at its output end. As the diameter of the fibre is very small, the light source has to be dimensionally compatible with fibre core. Light emitting diodes and laser diodes, which are very small in size, serve as the light sources. The electrical input signal is in general of digital form. It is converted into an optical signal by varying the current flowing through the light source. Hence, the intensity of the light emitted by the source is modulated with the input signal and the output will be in the form of light pulses. The light pulses constitute the signal that travels conductor photodiodes, which are very small in size, photo conductor converts the optical signal into electrical form. Thus, a basic optical fibre system consists of a LED/Laser diode, optical fibre cable and a semiconductor photodiode.

Optical fibre cables are designated in different ways for different applications. More propagations are provided to optical fibre by the cable which has the fibres and strength members inside an outer covering called a “Jacket”. We have two typical fibre cable, a signal fibre cable or a multifibre cable.

i. **Sigle Fibre Cable:** Around the fibre a tight buffer jacket of Hytrel is used (fig.3).

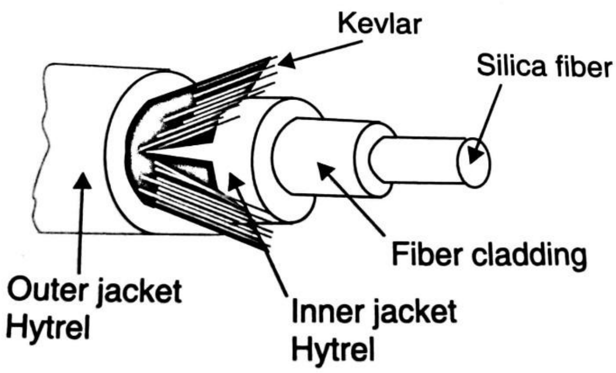


Figure : 3

The buffer Jacket protects the fibre from moisture and abrasion. A strength order to provide the necessary toughness and tensile strength. The strength member may be a steel wire, polymer film, nylon yarn or Kevlar yarn. Finally, the fibre cable is covered by a hytrel outer Jacket. Due to this arrangement fibre cable will not get damaged during bending, rolling, stretching or pulling and transport and installation processes, The single fibre cable is used for indoor applications.

ii. **Multifibre Cable:** A multifibre cable consists of a number of fibres in a signal Jacket. Each fibre carries light independently. The cross-sectional view of a typical multifibre cable is shown figure 4.

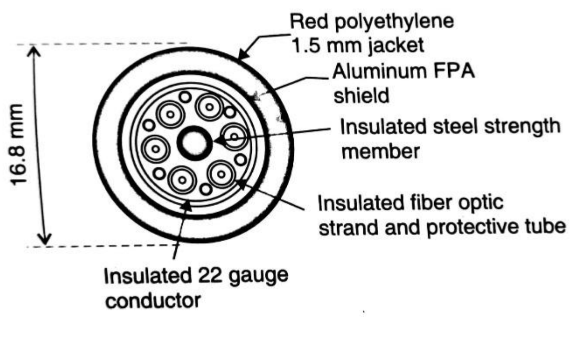


Figure: 4

It contains six insusteel cable at the centre for providing tensile strength. Each optical fibre strands consists of a core surrounded by a cladding, which is coated with insulating Jacket.

The fibres are thus individually buffered and strengthened. Six insulated copper wires are distributed in the space between the fibres. They are used for electrical transmission, if required. The assembly is then fitted with in a corrugated aluminum sheath, which acts as a shield. A polythene jacket is applied over the top.

3.5 PROPOGATION OF LIGHT IN OPTICAL FIBRE COMMUNICATIONS

The diameter of an optical fibre is very small and as such we cannot use bigger light sources for launching light beam into it. Light emitting diodes (LED's) and laser diodes are the optical sources used in fibre optics. Even in case of these small sized sources, a focusing lens has to be used to concentrate the beam on to the fibre core. Light propagates as an electromagnetic wave through an optical fibre. However, light propagates as an electromagnetic wave through an optical fibre can be well understood on the basis of ray model. According to ray model, light rays entering the fibre strike the core- clad interface at different angles, As the refraction index of the cladding is less than that of the core, majority of the rays undergo total inernal reflection at the interface and the angle of reflection is equal to angle of incidence in each case. Due to the cylindrical symmetry in the fibre structure, the rays reflected from an interface on one side of the fibre axis will suffer total internal reflections at the interface on the opposite side also. Thus, the rays travel forward through the fibre via a services of total internal reflections and image out from the exit end of the fibre (Fig. 5).

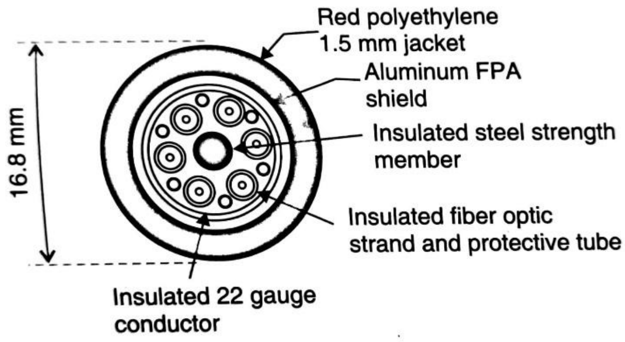


Figure : 5

Since each reflections is a total reflection, there is no loss of light energy and light confines itself within the core during the course of prorogation. Because of negligible loss during the total internal reflections, optical fibre can carry the light waves over very long distances. Thus, the optical fibre is essentially as a waveguide and is often called a light guide or light pipe. At the exist end of the fibre, the light is received by a photodetector.

Total internal reflection at the fibre wall can occur and light propagates down the fibre, only if the following two conditions are satisfy

1. The refractive index of the core n1 must be slightly greater than that of the cladding n2.
2. At the core- cladding interface (Fig. 6),

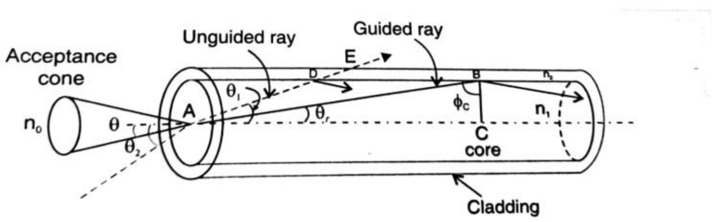


Figure : 6

The angle of incidence Ø between the ray and the normal to the interface must be greater than the critical angle Øc defined by

$$\sin\ \theta_c = \frac{n_2}{n_1}$$

It is noted that only those rays, that are incident at the core-cladding interface at angles greater than the critical angle will propogate through the fibre. Rays that are incident at smaller angles are refracted into the cladding and are lost.

3.6. ADVANTAGE OF OPTICAL FIBRE COMMUNICATIONS

Optical fibres have many advantageous features that are not found in conducting (copper) wires. Some of the important advantages are in follows.

1. Cheaper: Optical fibres are made of from silica (SiO2) which is one of the most abundant materials on the earth. The overall cost of a fibre optics communication is lower than that of an equivalent cable communication system.
2. Smaller in size, lighter in weight and flexible yet strong: The cross section an optical fibre is about a few hundred micros. Hence, the fibres are less bulky. Typically, A coasual cable weight about 1100 Kg/Km where as a fibre cable weights about 6 Kg/Km only. Optical fibres are quite flexible and strong.
3. Not hazardous: A wire communication link could accidently short circuit high voltage line and the sparking occurring there by coulignite combustible gases in the area leading to a great damage. Such accidents cannot occur with fibre links, because fibres are made of insulating materials.
4. Immunete EMI and RFI : In optical fibres information is carried by photon. Photons are electrically neutral and cannot be disturbed by high voltage fields, lighting etc. Therefore, fibres are immune to externally caused background noise generated through electromagnetic interference (EMI) and radio frequency interference (RFI).
5. No cross talk: The light waves propagating along the optical fibre are completely trapped within the fibre and cannot leak out, Further, light cannot couple into the fibre

from sides. Due to the sea features, possibility of cross talk is minimized. When optical fibre is used. Therefore, transmission is more secure and private.

6. Wider bandwidth : Optical fibres have ability to carry large amounts of information. A telephone cable composed of 900 pairs of wire can handle 10,000 calls, a 1 mm optical fibre can transmit 50,000 calls.
7. Low loss per unit length: The transmission loss per unit length of an optical fibre is about 4dB/Km. Therefore, longer cable runs between repeats are to be spaced at an interval about 2 km. In case of optical fibres, the interval can be a large as 100 km and above.

3.7 DISADVANTAGE OF OPTICAL FIBRE COMMUNICATIONS

Optical fibre communication system has few disadvantages as follows.

1. Low power: Light emitting sources are limited to low power. Although high power emitters are available to improve power supply. It would add extra cost.
2. Fragility: Optical fibre is rather fragile and more vulnerable to damage compared to copper wires. It is better not to twist or bend fibre optics cables too tightly.
3. Distance: The distance between the transmitter and receiver should keep short or repeaters are needed to boost the signal.

3.8 NUMERICAL APERATURE, ACCEPTATNCE ANGLE, V-PARAMETERS, MERIDIONAL AND SKEW RAYS ANALYSIS

Let us consider a step index optical fibre into which light is launched at one end as shown in fig. 7.

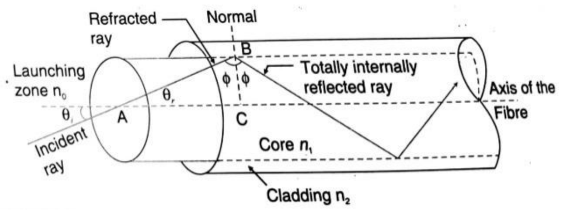


Figure: 7

Let the refractive index of core be n1 and refractive index of cladding is n2 (n2<n1). Let n0 be the refractive index of the medium from which light is launched into the fibre.

Assume that a light ray enters the fibre at an angle θi to the axis of the fibre. The ray refracts at an angle Φ . If Φ is greater than critical angle Φc, the ray undergoes total internal reflection at the interface, since n1>n2 as long the angle Φ is greater than Φc, light will stay in with in the fibre.

Applying snell's law at the launching face of the fibre, we get

Sinθi/Sinθr= n1/n0 1

If θi is increased beyond a limit, Φ will drop below the critical value Φc and ray escapes from the sidewalls of the fibre. The largest value of θi occurs when Φ=Φc

From Δ ABC , it is seen that

Sinθr= Sin(90-θ) = cosΦ

$$sin\theta_i = \frac{n_1}{n_2} cos\theta$$

When $\phi = \phi_c$ $sin[\theta_{i_{max}}] = \frac{n_1}{n_2} cos\phi_c$

But

$$\sin\phi_c = \frac{n_2}{n_1}$$

$$\cos\phi_c = \frac{\sqrt{n_1^2-n_2^2}}{n_1}$$

$$sin\left[\theta_i(max) = \frac{\sqrt{n_1^2 - n_2^2}}{n_0}\right]$$

Quite often the incident ray is launched from as medium, for which $n_0 = 1$.

Designating θ_i (max) = θ_0,

$$sin\theta_0 = \sqrt{n_1^2 - n_2^2}$$

$$\theta_0 = sin^{-1}\left[\sqrt{n_1^2 - n_2^2}\right]$$

The angle θ_0 is called the acceptance angle of the fibre. Acceptance angle is the maximum angle that a light ray can have relative to the axis of the fibre and propagate down the fibre. Thus only those rays that are incident on the face of the fibre making angles less than θ_0 will undergo repeated total internal reflections are reach other end of the fibre. Light incident at an angle beyond θ_0 refracts through the cladding and corresponding optical energy is lost.

Fractional Refractive Index Change

The fractional difference Δ between the refractive indices of the core and cladding is known as fractional refractive index change. It is repressed as

$$\Delta = \frac{n_1-n_2}{n_1}$$

This parameter is always positive because $n_1 > n_2$ for total internal reflection conduction. $\Delta << 1$ typically Δ ~- 0.01

Numerical Aperture

The main function of an optical fibre is accepted and transmit as much light from the source as possible. The light gathering ability a fibre depends on the numerical aperture. The acceptance angle and fractional refractive index change determine the numerical aperture of fibre.

The numerical aperature (NA) is defined a sine of the acceptance angle, Thus

NA = Sinθ_0

Where θ_0 is acceptance angle

$$Sin\,\theta_0 = \sqrt{n_1 - n_2}$$

$$NA = \sqrt{n_1^2 - n_2^2}$$

$$n_1^2 - n_2^2 = (n_1 + n_2)\,(n_1 - n_2) = \left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1-n_2}{n_1}\right)2n_1$$

Approximating $\left(\frac{n_1+n_2}{n_1}\right)2n_1$, we can express above relation as $(n_1^2 - n_2^2) = 2n_1^2\,\Delta$. It gives

$$NA = \sqrt{2n_1^2\,\Delta}$$

$$NA = n_1\sqrt{2\Delta}$$

Numerical aperature determines the light gathering ability of the fibre. It is a measure of the amount of light that can be accepted by a fibre. It is clear that N.A. is depedent only on the refractive indices of the core and cladding materials and does not depend on the physical dimesions of the fibre. The value of NA ranges from 0.13 to 0.50. A large NA implies that a fibre will accept large amount of light from the source.

V-Number (V- parameter)

When a narrow beam of monochromatic light launched on the front end of a step index fibre, at an angle less than the acceptance angle of the fibre. Let the wavelength of the light be$\lambda 0$ and the diameter of the fibre be d. From the ray concept it is clear that all the rays contained in the beam propagate along the fibre, such that there can be infinite modes of propagation. However, in practice only a limited number of modes of propagation are possible in an optical fibre. Thus, is due to fact that phase changes occur as the light waves travel forward. The phase shift takes place due to two reasons

I. Due to optical path length traversed and

II. Due to total internal reflection at the core – cladding interface.

III. When a wave travels a distance l in a medium of refractive index n1, it undergoes a phase

Change δ_1 given by

$$\delta_1 = kn_1 l = \frac{2\pi\ l\ n_1}{\lambda}$$

Where k is propogation constant.

i. Whenever, a wave with component normal to the reflecting surface undergoes total internal reflection, the phase shift δ_2 is given by

$$\delta_2 = 2\ tan^{-1} = \frac{\sqrt{n_1^2 cos^2\phi - n_2^2}}{n_1 \sin\phi}$$

Combining above equation we get,

$$\delta = \frac{4d\pi\ n_1 \cos\phi}{\lambda_0} - 2\delta_2$$

Total phase shift must be equal to an integral multiple of 2π radians. Thus

$$\frac{4d\pi\ n_1 \cos\phi}{\lambda_0} - 2\delta_2 = 2\pi\ m$$

$$m = \frac{2d\ n_1 \cos\phi_m}{\lambda_0} - \frac{\delta_2}{\pi}$$

Where m is an integer that allowed ray angles for propogation of the wave and Φm is the value of Φ corresponding to a particular value of m. In order to sustain total internal reflection

$$\sin \phi_m \geq \frac{n_2}{n_1}$$

$$\cos \phi_m \leq \frac{\sqrt{n_1^2 - n_2^2}}{n_1}$$

$$m \leq \frac{2d\sqrt{n_1^2 - n_2^2}}{\lambda_0} - \frac{\delta_2}{\pi}$$

$$m \leq \frac{2V}{\pi} - \frac{\delta_2}{\pi}$$

Where V is given by

$$V = \frac{\pi\, d}{\lambda_0}\sqrt{n_1^2 - n_2^2}$$

V – number is more generally called normalized frequency of the fibre. Each mode has a definite value of V- number below which the mode is cut off,

$$V = \frac{\pi\, d}{\lambda_0}\,(NA)$$

$$V = \frac{\pi\, d}{\lambda_0} n_1\sqrt{2\Delta}$$

The maximum number of modes Nm supported by an SI fibre is given by

$$N_m = \frac{1}{2}V^2$$

Thus for v= 10, Nm= 50 . When the normalized frequency v is less than 2.405, The fibre can support only one mode, which propagates along the axial length of the fibre end the fibre becomes a single mode fibre. It means that for single mode transmission in a MMF, V must be less than 2.405. The wavelength at which fibre becomes single mode is called cut off wavelength λ_c of the fibre.

$\lambda_c = \frac{\pi d}{2.405}\ (NA)$

It is seen that single mode property can be realized in a multimode fibre by decreasing the core diameter and/or decreasing Δ such that V< 2.405.

In case of graded index fibre for large value of v

Nm $\cong v^2/4$

Meridinal and Skew Ray analysis

The rays that propagate through an optical fibre can be classified into two categories:

I. Meridional ray

II. Skew rays

1. Meridinal rays : A ray that propogates through the fibre undergoing total internal reflection called merdional rays. It passes through the longitudinal axis of the fibre core (Fig)8a). The propogation of meridional rays is possible only in the TM or Te modes.

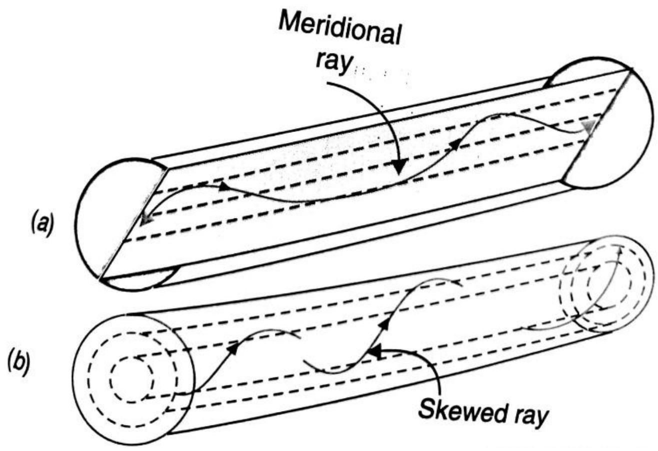


Figure: 8

2. Skew ray : The ray that describes angular helical path along the fibre is called a skew ray (Fig)8b). These rays are propogated in either hybrid EH or HE modes. Some of

these modes produce losses through leakage of radiation. In real situations, the skew rays constitute a substantial portion of the total number of guided rays. They tend to propagate only in the annular region near the outer surface of the core as medium. However, they are complementary to the meridional rays and increase the light gathering capacity of the fibre.

3.9 CLASSIFICATION OF FIBRES

Optical fibres are differently classified into various types based on different parameters.

Classification based on Reference Index Profile

Refractive index profile of an optical fibre is a plot of refractive index on one axis and distance of the core on the other axis. Optical fibres are classified into the following two categories on the basis of refractive index profile.

1) Step index fibre

2) Graded Index (GRIN) Fibres. (Fig.9)

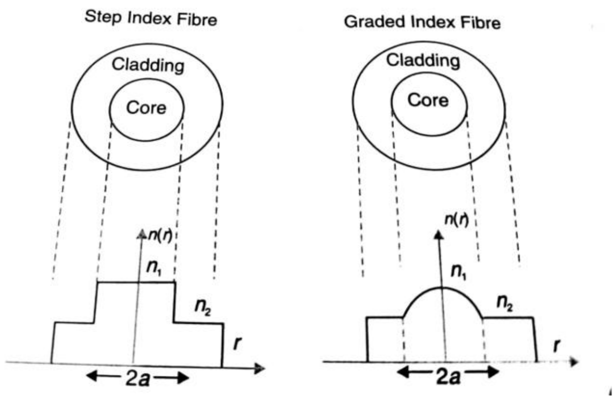


Figure : 9

Step index refers to the fact that refractive index of the core is constant along the radial direction and abruptly falls to a lower value at the cladding to core boundary. (Fig.9a)

In case of Graded index fibre (GRIN), the refractive index of the core is not constant but varies smooth over the diameter of the core. (Fig 9b). It has maximum value at centre and decreases gradually towards the outer edge of the core. At the core- cladding interface the refractive index of the core matches with the refractive index of the cladding. The refractive index of the cladding is constant.

Classification based on modes of light propogation

On basis of modes of light propogation, optical fibres are classified into two categories as.

I. Single mode fibres (SMF)

II. Multimode fibres (MMF)

A single mode fibre has a smaller core diameter and can support only one mode of propogation on the other hand, a multimode fibre (MMF) has a larger core diameter and supports a number of modes.

Thus on the whole, the optical fibres are classified into three types.

- Single mode step index fibre (SMF)
- Multimode step index fibre (MMF)
- Graded index (multimode) (GRIN) Fibre

Classification based on Materials

On the basis of materials used for core and cladding, optical fibres are classified in three categories

1. Glass/glass fibres (Glass core with Glass cladding)
2. PCS (Polimeter clad silica)

(Polymer as cladding and silica as core).

Three types of Fibres

Here we study the detailed structure and characteristics of the three types of optical fibres.

Single Mode Step Index Fibre :

Structure : A single mode step index fibre has a very fine thin core of diameter of 8 Km to 12 Km (see fig.9c). It is usually made of germanium doped silicon. The core is surrounded by a thick cladding of lower refractive index. The cladding is composed of silica lightly doped with phosphorous oxide. The external diameter of the cladding is of the order of 125Km. The fibre is surrounded by an opaque protective sheath. The refractive index of the fibre changes abruptly at the core- cladding boundary, as shown in fig. 10(a). The variation of the refractive index of a step index fibre as a function of radial distance can be mathematically represented as

$n(r) = n_1$ ($r<a$ inside core)

$= n_2$ ($r>a$ in cladding)

Propogation of light in SMF

Light travels SMF along a single path that is along the axis (Fig 10b). It is the zero order mode that is supported by a SMF. Both Δ and NA are very small for single mode fibres. This relatively small value is obtained by reducing the fibre radius and by making Δ, the relative refractive index change to be small. Therefore, light coupling into the fibre becomes difficult. Costly laser diodes are needed to launch light into the SMF.

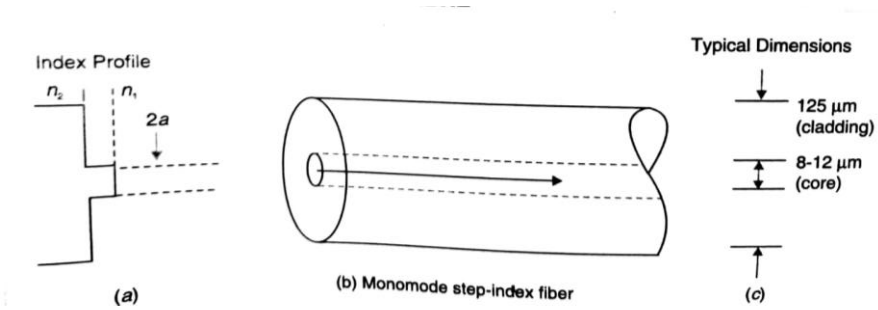


Figure : 10

Multimode Step Index Fibre

A multimode step index fibre is very much similar to the single mode step index fibre except that its core is of larger diameter. The core diameter is of the order of 50 to 100 Km, which is very large compared to the wavelength of light. The external diameter of cladding is about 150 to 250 Km.

Propogation of Light in MMF

Multimode step index fibres allow finite number of guided modes. The direction of polarization, alignment of electric and magnetic fields will be different in rays of different modes. In other word many zig-zag paths are longer. Because of this difference, the lower order modes reach the end of the fibre earlier while the light order modes reach after some time delay (Fig 11b)

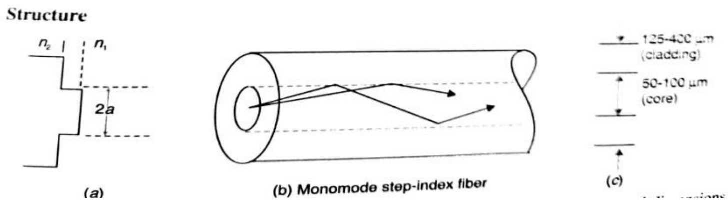


Figure: 11

Graded Index (GRIN) Fibre

A graded index fibre is a multimode fibre with a core, consisting of concentric lasers of different refractive indices. Therefore, the refractive index of the core varies with distance from the fibre axis. It has a high value at the centre and falls of with increasing radial distance from the axis. It has high value at centre and falls of with increasing radial distance from the axis. It has high value at centre and fall of with increasing radial distance from the axis. A typical structure and its index profile are shown in fig. 12(a).

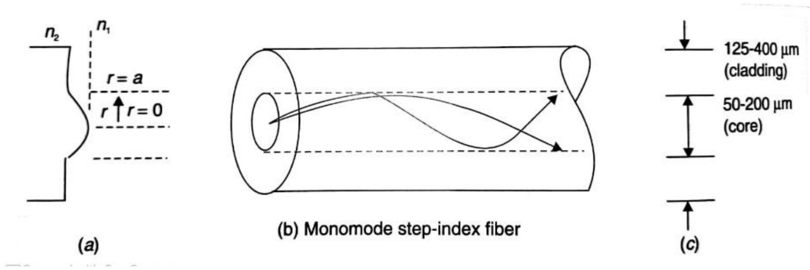


Figure: 12

Such a profile causes a periodic focusing the graded index fibre is about the same as the step index fibre.

The variation of the refractive index of the core with radius measured from the centre is given by

$$n(r) = \begin{Bmatrix} n_1\sqrt{1 - \left[2\Delta\left(\frac{r}{a}\right)^{\alpha}\right]}, & r < a \; inside \; core \\ n_2, & r > a \; in \; cladding \end{Bmatrix}$$

Where $n_1$is maximum refractive index at the core axis, a the core radius, and ∝ is the grade profile index number which varies from 1 to ∞ .When∝=2, the index profile is parabolic and preferred for different applications.

Propogation of light

As a light ray goes from a region of higher refractive index to a region of refractive index. It is bent away from the normal. The process continuous till the condition for total internal reflection is met. Then the ray travels back towards the core axis, again being continuously refracted (Fig. 13).

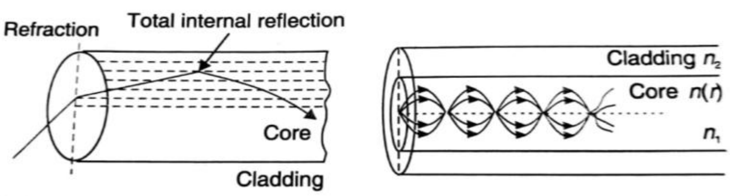


Figure: 13

The turning around may take place even before reaching the core- cladding interface. Thus, continuous refraction is followed by total internal reflection and again continuous refraction towards the axis. In the graded index fibre, rays making large angles with the axis traverse longer path but the travel in a region of lower refractive index and hence at a higher speed of propagation consequently all rays travelling through the fibre, irrespective of their modes of travel, will have almost the same optical path length and reach the output end of the fibre at the same time.

In case of GRIN fibres, the acceptance angle and numerical aperture decrease with radial distance from the axis. The numerical aperture of a graded index fibre is given by

$$\text{NA}= \sqrt{n^2(r) - n_2^2} = n_1(2\Delta)^{\frac{1}{2}}\sqrt{1-\left(\frac{r}{a}\right)^2}$$

$$= n_1\sqrt{2\Delta\left[1-\left(\frac{r}{a}\right)^2\right]}$$

3.10 ATTENUATION AND DISPERSION IN OPTICAL FIBRES (ATTENUATION LOSS, DISPERSION)

As a light signal propagates through a fibre, it suffers loss of amplitude and charge in shape. The loss of amplitude is referred to as attenuation and change in shape as distortion.

Attenuation : When an optical signal propagates through a fibre, its power decrease exponentially with distance. The loss of optical power as light travels down a fibre is known as attenuation. The attenuation of optical signal is defined as the ratio of optical output power from a fibre of length L down the fibre is given power $P_0 = P_i\, e^{-\alpha L}$

Where $\propto$ is called the fibre attenuation coefficient expressed in units of Km^{-1} taking logarithims on both sides of the above equations, we obtain

$$\propto = \frac{1}{L}\ln\frac{p_i}{p_o}$$

In units at dB/Km,$\propto$ is difiened through the equation

$$\propto_{dB/Km} = \frac{10}{11} \log \frac{p_i}{p_o}$$

In case of ideal fibre $P_0 = P_i, \propto = 0$

Types of Attenuation

There are several loss mechanisms responsible for attenuation in optical fibres. They are broadly divided into two categories. Intrinsic and Extrinsic attenuation.

A. <u>Intrinsic Attenuation:</u> Intrinsic attenuation results from materials inherent to the fibre. It is caused by impurities present in the glass. During manufacturing, there is no way to eliminate all impurities. When a light signal hits an impurity in the fibre, either it is scattered or it is absorbed.

   Intrinsic attenuation can be further divided by in the two parts.

I. Material absorptions
II. Rayleigh scattering

1. Material Absorption

Material absorption occurs as a result of the imperfection and impurities in the fiber and accounts for 3.5% of fiber attenuation. The natural impurities in the glass absorb light signal and convert it into vibrational energy or some other form of energy. Hydroxyl radical ions (OH), and transition metals such as copper, nickel, chromium, vanadium and magnese have electronic absorption in and near visible part of the spectrum. Their presence causes heavy losses.

Unlike scattering, absorption can be limited by controlling the number of impurities during the manufacture process.

2. Rayleigh Scattering

Rayleigh scattering accounts for majority (about 96%) of attenuation in optical fibre. The local microscopic density variations in glass cause local variations in refractive index. The variations which are inherent in the manufacturing process and cannot be eliminated, acts as obstructions and scatter light in all directions (Fig. 14). This is known as Rayleigh scattering. The Rayleigh scattering loss greatly depends on the wavelength. It varies as 1/λ4 and becomes important at lower wavelengths.

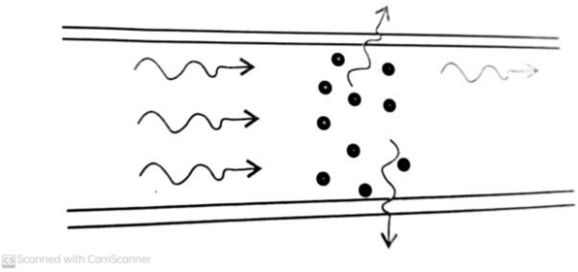


Figure: 14

Rayleigh scattering sets a lower limit on the wavelengths that can be transmitted by a glass fibre at 0.8km, below which the scattering loss is very high. Any wavelength that is below 800nm is not suitable for optical communication because attenuation due to Rayleigh scattering is high. At the same time propagation above 1700 nm is not possible due to high losses resulting from infrared absorption.

For better performance, the choice of wavelength must be based on minimizing loss and minimizing dispersion. It is seen from the attenuated around a particular band of optical wavelengths. The band of wavelengths at which the attenuation is minimum is called optical window. There are three principal windows. These corresponds toward length regions in which attenuation is low and matched to the capability of a transmitter to generate light efficiently and receiver to carry out detection.

λ (nm)	Approx loss (dB/Km)
820- 880	2.2

1200-1320	0.6
1500-1610	0.2

From above data it is seen that the range 1550- 1610 is most favorable.

B. Extrinsic Attenuation or Bending losses

Extrinsic attenuation is caused by two external mechanisms. Micro bending and Micro bending .

Microband losses

A microband is a large- scale bend that is visible. When a fibre is bent through a large angle, strain is placed on the fiber along the region that is bent. The bending strain will affect the refractive index and the critical angle of the light ray in that specific area. As a result, light travelling in the core can refract out and loss occurs. (Fig.15)

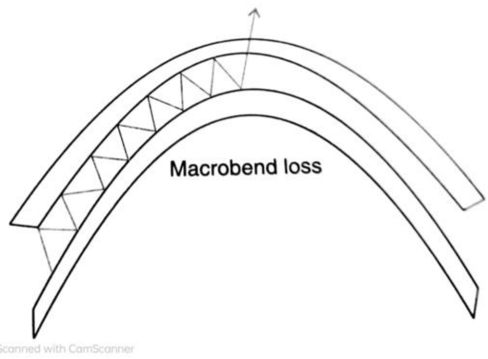


Figure: 15

Microbend

Microband is a small-scale distortion. It is localized and generally indicative of pressure on the fibre. Micro bending might be related to temperature, tensile stress, or crushing force. Micro bending is caused by imperfections in the cylindrical geometry of fibre during the manufacturing process or installation processes.

The bend may not be clearly visible upon inspection structural variation in the fibre or fibre deformation cause radiation of light away from the fibre. (Fig. 16)

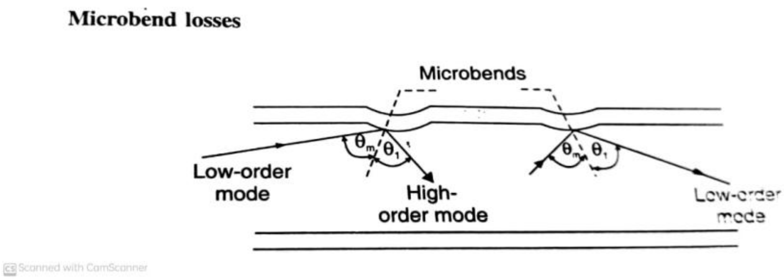


Figure: 16

Dispersion in optical Fibre

Optical Fibre dispersion describes the process of how an input signal broadens/ spreads out as it propagates/ travel down the fiber. There are three different types of dispersion mechanisms that determine distortion of the signal in an optical fibre. They are

i. Intermodal dispersion and

ii. Intramodal dispersion

Intramodal dispersion is again divided into the following two types:

a) Material dispersion

b) Waveguide dispersion

I. Intermodal dispersion

Intermodal dispersion occurs as a result of the differences in the group velocities of the modes. For example, let us consider the propagation of a pulse through a multimode fibre. The lower order modes (ras reflected at larger angles) travel a greater distance than the higher order modes (lower angle rays). The path length along the axis of the fibre is shorter while that along zigzag path is longer. Because of this difference, the lower order modes reach the end of the fibre earlier while the high order modes reach after sometime delay. As a result, light pulses broaden as they travel down the fibre, causing signal distortion, the output pulses no longer resemble the input pulses.

This type of distortion is known as intermodal or simply modal dispersion.

Total time delay due to modal dispersion in step index fibre is given by

$$\Delta t = \frac{L}{2n_2c}(NA)$$

Where L- length of fibre travel by high

c- speed of light

n_2- refractive index of the cladding

NA- Numerical aperature of fibre

$From$ equation it is clear that time delay is proportional to square of the value of N.A. Thus for low modal dispersion N A of fibre should be small.

II. Intramodal Dispersion

Intramodal dispersion is spreads of light pulse with in a single mode. The two main causes of intramodal dispersion are

a) Material Dispersion

b) Waveguide Dispersion

a. Material Dispersion

Glass is a dispersive medium. A light pulse is a wave packet, composed of a group of components of different wavelengths. The different wavelength components will propagate at different speeds along the fibre (Fig 17). The short wavelength components travel slower than long wavelength components, eventually causing the light pulse to broaden. This type of dispersion is known as material dispersion. It is often called th chromatic dispersion.

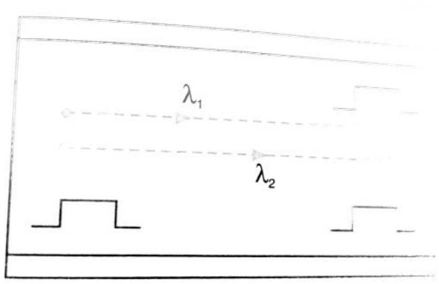


Figure: 17

Expression for material dispersion

$D_{mat}(\lambda) = \frac{\lambda}{c} d n^2 / d\lambda^2$

From equation it is clear that the material dispersion can be reduced either by choosing sources with narrows spectral range or by operating at longer wavelengths

b. Wave-guide Dispersion

Waveguide dispersion arises from the guiding properties of the fibre. The group velocities of modes depend on the wavelength. Hence the effective refractive index for any mode varies with wavelength. It is equivalent to the angles between the ray and the fibre axis varying with wavelength which consequently leads to a variation in the transmission times for the rays and hence dispersion.

Total Dispersion

All the above three dispersions contribute pulse spreading during signal transmission through an optical fibre. The total dispersion introduced by an optical fibre is given by the root mean square value of all the three dispersions.

Thus

$$(\Delta t)_T = \sqrt{(\Delta t)^2_{int\,er\,mod\,al} + (\Delta t)^2_{mat} + (\Delta t)^2_{wg}}$$

3.11 QUALITATIVE DISCUSSION OF COUPLER, SPLICES AND CONNECTORS

Coupler

A fibre optic coupler is an optical device capable of connecting one or maximum fibre ends in order to allow the transmission of light waves in multiple paths. The devices capable of combining two or more inputs into single output and also dividing a single input into two or more outputs. Compared to a splice or connector, the signal can be more attenuated by fiber optic couplers, as the input signal can be divided amongst the output.

I. <u>Splices and Connector:</u> (optics by Kar)

The most important factor in the installations any fibre optic system is the interconnection of fibres with minimum possible loss. The interconnections are needed at the optical sources in the transmitter at the photodetector in the receive and at intermediate points within a cable where two fibers are joined together. The particular technique for joining two fibers depression.

Whether a permanent bond or easily remountable (or demountable) connections is desired. The permanent bond is known as a splice, where as a demountable joint is referred as a connector. Every Joining is subject to certain conditions which become responsible for causing various types of optical power loss. The amount of loss depends on different parameters such as the input power distribution at the joint, the geometric and wavelength characteristics of two fibres which are coupled, various types of misalignments between the fiber ends and joints and the qualities of the fiber and its faces.

The optical fiber power can be coupled from one fiber to another which is controlled by the number of modes that can propagate in each fiber. For example, a fiber in which 500 modes can propagate is connected to another fiber in which 400 modes can propagate then at best 80% of the optical power from the first fiber can be couples to the second one provided all modes are equally excited.

3.12 SUMMARY

Fiber optics is the technology used to transmit information as pulses of light through strands of fiber made of glass or plastic over long distances.

Optical fibers are about the diameter of a strand of human hair and when bundled into a fiber-optic cable, they’re capable of transmitting more data over longer distances and faster than other mediums. It is this technology that provides homes and businesses with fiber-optic internet, phone and TV services.

3.13 TERMINAL QUESTION

1. What is Optical Fibre.
2. Write down the Principle of Optical Fibre.
3. What is attenuation in Optical Fibre.
4. What is an optical fibre? What is the principle involved in its working?
5. Explain the following terms:
   (a) Critical angle
   (b) Acceptance cone
   (c) Numerical aperture

3.14 ANSWER & TERMINAL QUESTIONS

1. Section 3.3
2. Section 3.4
3. Section 3.10

4. Section 3.3/3.4

5. (a) Section 3.8

   (b) Section 3.8

   (c) Section 3.8

3.15 SUGGESGTED READING

1. A Textbook of Optics Dr. N. Subrahmanyam Brij Lal & Dr. M. N. Avadhanulu

2. Optics – Ajoy Ghatak

3. Optics – Principles and Application, K. K. Sharma, Academic Press.

4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.

5. Advanced Geometrical Optics – Psang Dain Lin.

DCEPHS-105

OPTICS

Uttar Pradesh Rajarshi Tandon

Open University

Bachelor of Science

Block

2 LIGHT and POLARIZATION

Course Design Committee

Prof. Ashutosh Gupta **Chairman**
Director, School of Science, UPRTOU, Prayagraj
Prof. A. K. Rai **Member**
Professor, Dept. of Physics, University of Allahabad, Prayagraj
Gorakhpur
Prof. Ram Kripal **Member**
Professor, Dept. of Physics, University of Allahabad, Prayagraj
Dr. Anjani Kumar Singh **Member**
Associate Professor (Retd.), E.C.C University of Allahabad, Prayagraj

Course Preparation Committee

Dr. Akhilesh Kumar Singh **Author**
Associate Prof. **Unit- 02-03**
Dept. of Physics
Pt. DDUGG, PG college Rajaji Puram, Lucknow

Dr. Ratana Kumar **Author**
Asst. Prof. **Unit- 06-10**
Dept. of Physics
RIET,college , Phagwara , Punjab

Dr. Arvind Kumar Mishra **Author**
Assistant Professor , Physics, **Unit- 01, 04 & 05**
School of Science, UPRTOU, Prayagraj

Prof. Nitya Kaushal Kishore **Editor**
Ex. HOD, Dept. of Physics, **Block -I,II & III**
S.K.M.U. University, **Unit- 01 to 10**
Dumka, Jharkhand.

Dr. Arvind Kumar Mishra **Coordinator**
Assistant Professor , Physics,
School of Science, UPRTOU, Prayagraj

DCEPHS – 105 : Optics
ISBN-

Printed By: Chandrakala Universal Pvt. 42/7 Jawahar Lal Nehru Road, Prayagraj.

UNIT- 4 NATURE OF LIGHT

STRUCTURE:

4.1 INTRODUCTION

Light is a form of radiant energy, that is, energy emitted by excited atoms or molecules which can cause the sensation of vision in a normal human eye.

The branch of optics that deals with the production, emission and propagation of light, its nature and the study of the phenomena of interference, diffraction and polarization is called physical optics. The basic principles regarding the nature of light were formulated in the latter half of the seventeenth century. Until about this time, the general belief was that light consisted of a stream of particles called corpuscles. These corpuscles were given out by a light source (an electric lamp, a candle, sun etc.) and they travelled in straight lines with large velocities. The originator of the emission or corpuscular theory was Sir Isaac Newton. According to this theory, a luminous body continuously emits tiny, light and elastic particles called corpuscles in all directions. These particles or corpuscles are so small that they can readily travel through the interstices of the particles of matter with the velocity of light and they possess the property medium. When these particles fall on the retina of the eye, they produce the sensation of vision. On the basis of this theory, phenomena like rectilinear propagation, reflection and refraction could be accounted for, satisfactorily. Since the particles are emitted with high speed from a luminous body, they, in the absence of other forces, travel in straight lines according to Newton's second law of motion. This explains rectilinear propagation of light.

4.2 OBJECTIVES

After studying this unit, student should be able to:

- Describe nature of Light.
- To understand the concept of Newton's Corpuscular Theory.
- Describe Huygens' Wave Theory.
- Explain Laws of Reflection and Refraction.
- To understand the Concept of Fermat Principles.

4.3 WHAT IS LIGHT?

Light:

1. Light is a type of electromagnetic radiation that allows the human eye to see or makes objects visible.
2. It is also defined as visible radiation to the human eye. Photons, which are tiny packets of energy, are found in light.
3. Light always moves in a straight line.
4. Light travels at a faster rate than sound. The speed of light is 3×10^8 m/s.
5. Light is a transverse wave that travels without the use of a medium.
6. Light does not require a physical medium to travel. That is, it can also move through a vacuum.

4.4 WHAT IS NATURE OF LIGHT?

Light is a form of energy (optical energy) which helps us in seeing objects by its presence.

Theories about nature of light:

1. **Particle nature of light (Newton's corpuscular theory):**

   According to Newton light travels in space with a great speed as a stream of very small particles called corpuscles.

   This theory was failed to explain interference of light and diffraction of light. So, wave theory of light was discovered.

2. **Wave nature of light:**

   Light waves are electromagnetic waves so there is no need of medium for the propagation of these waves. They can travel in vacuum also. The speed of these waves in air or in vacuum is maximum i.e., 3×10^{8}m/s.

   Photoelectric effect was not explained with the help of wave theory, so Plank gave a new theory which was known as quantum theory of light.

3. **Quantum theory of light:**

When light falls on the surface of metals like calcium, potassium etc., electrons are given out. These electrons are called 'photo-electrons' and phenomenon is called 'photo-electric effect'.

This was explained by Einstein. According to plank light consisted of packets or quanta of energy called photons. The rest mass of photon is zero. Each quanta carries energy

$E = h\nu$.

$h \rightarrow$ Planck's constant = 6.6×10^{-34} J-s.

$\nu \rightarrow$ Frequency of light

Some phenomenon like interference of light, diffraction of light is explained with the help of wave theory but wave theory was failed to explain the photo electric effect of light. It was explained with the help of quantum theory. So, light has dual nature.

1. Wave nature

2. Particle nature

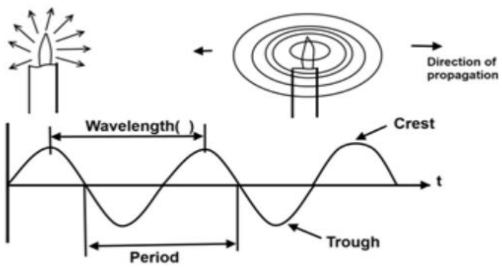


Figure : 1

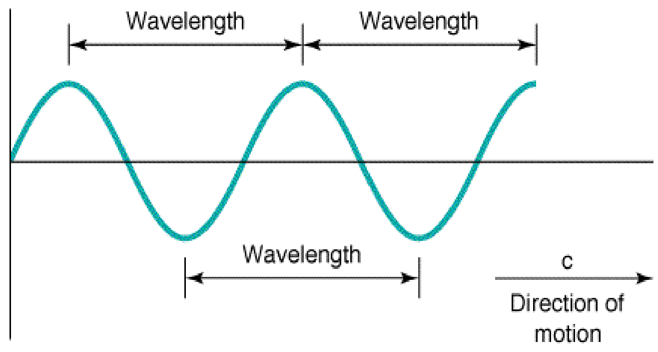


Figure : 2

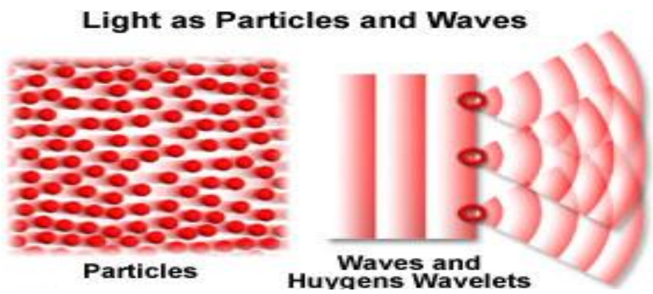


Figure : 3

	Phenomenon	Can be explained in terms of waves.	Can be explained in terms of particles.
2	Reflection	✓	✓
3	Refraction	✓	✓
4	Interference	✓	⊗
5	Diffraction	✓	⊗
6	Polarization	✓	⊗
7	Photoelectric effect	⊗	✓

4.5 NEWTON'S CORPUSCULAR THEORY

Sir Issac Newton proposed the corpuscular theory to explain the properties of light. According to this theory, **light is made from small and extremely light particles called corpuscles**. When these corpuscles travelling in straight line hit the retina of the eye, it produces the sensation of vision.

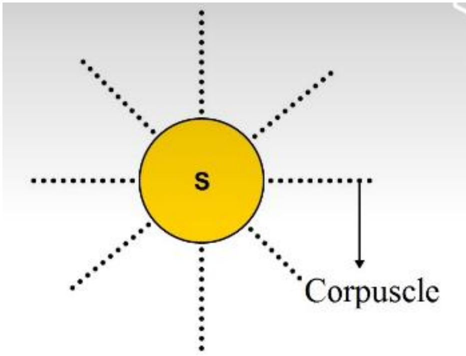


Figure : 4

4.5.1 Statement of Newton's Corpuscular Theory

Newton's corpuscular theory was based on postulates as follows:

1. Newton's proposed that a source of light emits many minute, elastic, rigid and massless particles called corpuscles.

2. These particles travel through a transparent medium at very high speed in all direction along a straight line.

3. These corpuscles enter our eyes and produce the sensation of vision.

4. Due to different sizes of the corpuscles, they produce different colors.

5. These light particles are repelled by a reflecting surface and attracted by transparent materials.

4.5.2 Merits of Newton’s Corpuscular Theory

1. It explains the rectilinear propagation of light.

2. It could explain reflection and refraction of light **separately**.

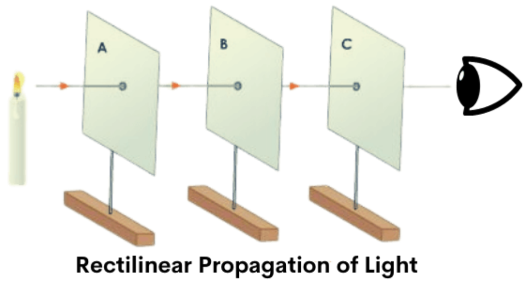


Figure : 5

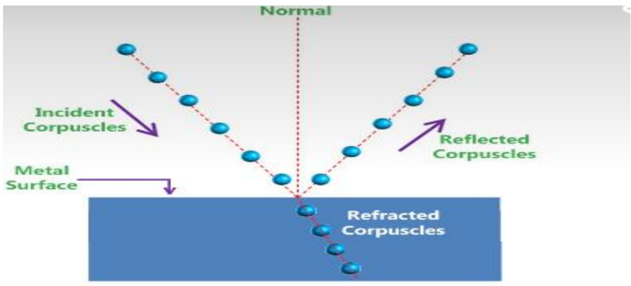


Figure : 6

4.5.3 Demerits of Newton’s Corpuscular Theory

1. Newton’s corpuscular theory fails to explain **simultaneous phenomenon of partial reflection and refraction** on the surface of transparent medium such as glass or water.

2. The corpuscular theory fails to explain optical phenomena such as interference, diffraction, polarization etc.

3. According to this theory, velocity of light is larger in the denser medium than in the rarer medium, experimentally it is proved wrong ($v_a < v_d$).

4. As the particles are emitted from the source, mass of the source of light should decrease but experiment proved that mass of the source of light is constant.

4.6. MAXWELL'S ELECTROMAGNETIC THEORY

According to this theory light is electromagnetic waves. Experimentally it is observed that velocity of light is equal to velocity of electromagnetic waves. They travel even in vacuum.

SAQ:

1. Newton's corpuscular theory could explain correctly the phenomenon of:

   (a) interference of light

   (b) diffraction of light

   (c) rectilinear propagation of light

   (d) simultaneous reflection and refraction of light.

Q2. What was Maxwell's concept of light?

Q3. Explain in brief the Newton's Corpuscular theory of light.

Q4. What are the draw backs of Newton's Corpuscular theory of light.

Q5. What is the photon model or quantum hypothesis of light?

Q6. Explain Max Plank's quantum theory of light.

4.7 MAX PLANKS QUANTUM THEORY

According to this theory light is propagated in the form of packets of energy called quanta. Each quanta of light also called photon and it has energy

$E = h\,v$

Where, v = Frequency of light

H = Plank's constant

4.8 HUYGENS' WAVE THEORY

In the late 17th century, scientists were embroiled in a debate about the fundamental nature of light – whether it was a wave or a particle. Sir Isaac Newton was a strong advocated of the particle nature of light. But, the Dutch physicist, Christiaan Huygens believed that light was made up of waves vibrating up and down perpendicular to the direction of the wave propagation, and therefore formulated a way of visualizing wave propagation. This became known as '<u>Huygens' Principle</u>'.

The wave theory of light proposed by Christian Huygens has stood the test of time, and today, it is considered the backbones of optics.

History of the Wave Theory of Light

Light always piqued the curiosity of thinkers and scientists. But it wasn't until the late 17th century that scientists began to comprehend the properties of light. Sir Isaac Newton proposed that light was made of tiny particles known as the photons while Christian Huygens believed that light was made of waves propagating perpendicular to the direction of its movement.

In 1678, Huygens proposed that **every point that a luminous disturbance meets turns into a source of the spherical wave itself. The sum of the secondary waves, which are the result of the disturbance, determines what form the new wave will take**. This theory of light is known as the 'Huygens' Principle'.

Using the above-stated principle, Huygens was successful in deriving the laws of reflection and refraction of light. He was also successful in explaining the linear and spherical propagation of light using this theory. However, he wasn't able to explain the diffraction effects of light. Later, in 1803, the experiment conducted by Thomas Young on the interference of light proved the Huygens wave theory of light to be correct. Later in 1815, Fresnel provided mathematical equations for Young's experiment.

Max Planck proposed that light is made of finite packets of energy known as a light quantum and it depends on the frequency and velocity of light.

Later, in 1905, Einstein proposed that light possessed the characteristics of both particle and wave. He suggested that light is made of small particles called photons. Quantum mechanics gave proof of the dual nature of light.

4.8.1 Postulates of Huygens' Wave Theory

Postulates on which Huygens' wave theory are given as follows:

a. The source of light emits in the form of waves.

b. Light waves are like a sound wave, which are longitudinal in nature.

c. Light waves move with constant speed in a homogeneous medium.

d. Different colors of light are due to different wavelengths of light waves.

e. When light waves enter in our eyes, we feel the sensation of vision.

f. Light waves travel through vacuum due to presence of a hypothetical medium called as luminiferous ether.

4.8.2 Merits of Huygens' Wave Theory

a. Wave theory of light is helpful to explain phenomena such as reflection, refraction, interference and diffraction.

b. The phenomenon of Partial reflection and refraction of light can be satisfactorily explained using the wave theory of light.

c. As per the wave theory of light, velocity of light in optically denser medium is less than the velocity of light in a rarer medium, which is correct $(v_a > v_d)$.

4.8.3 Demerits of Huygens' Wave Theory

a. Wave theory of light assumed the presence of hypothetical ether medium but experiment proved that there is no ether or drag.

b. Rectilinear propagation of light is not explained by this wave theory.

c. Wave theory of light could not explain phenomena such as Compton effect and polarization of light.

d. Wave theory of light could not explain bending of wave through an obstacle.

e. Wave theory of light assumed that light waves are longitudinal in nature but experiment proved that they are electromagnetic transverse waves.

SAQ:

Q7. Give brief account of Huygens' wave theory of light. States its merits and demerits.

Q8. Huygens' wave theory of light could not explain:

(a) reflection

(b) refraction

(c) interference

(d) photoelectric effect

Q9. Huygens' original theory of light assumed that light propagates in the form of:

(a) transverse mechanical waves.

(b) longitudinal mechanical waves.

(c) transverse electromagnetic waves.

(d) minute elastic particles.

Q10. The phenomenon of diffraction and refraction indicates that light is having:

(a) particle nature

(b) wave nature

(c) both particle and wave nature

(d) neither particle nor wave nature

Q11 Two points, equidistant from a point source of light, ae situated at diametrically opposite positions in an isotropic medium. The phase difference between the light waves passing through the two points is:

(a) zero

(b) π rad

(c) $\pi/2$ rad

(d) 2π rad

4.9 IMPORTANT TERMINOLOGY

There are same Important Terminology based on wave theory :

4.9.1 Wave Surface

When a point source of light 'S' is situated in air then its waves travel. In all possible directions. If 'c' is the velocity of light in air then each wave covers a distance 'ct' in time t and reaches the surface of a sphere.

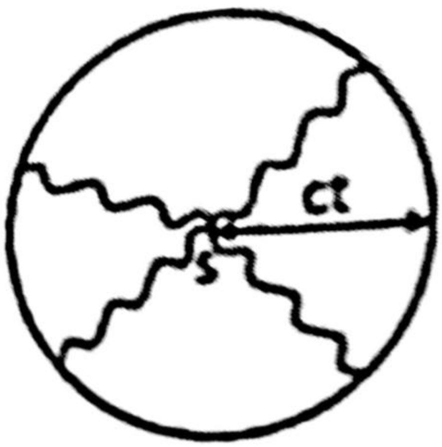


Figure : 7

4.9.2 Wavefront

"The locus of all points of the medium at which the waves reach simultaneously such that all points are in the same phase is called a wavefront ".

There are three type of wavefront :

I. Spherical wavefront

II. Plane wavefront

III. Cylindrical wavefornt

4.9.3 Spherical Wavefront

A wavefront in the form of spherical surfaces is called spherical wavefront.

It is obtained from a point source of light up to a finite distance.

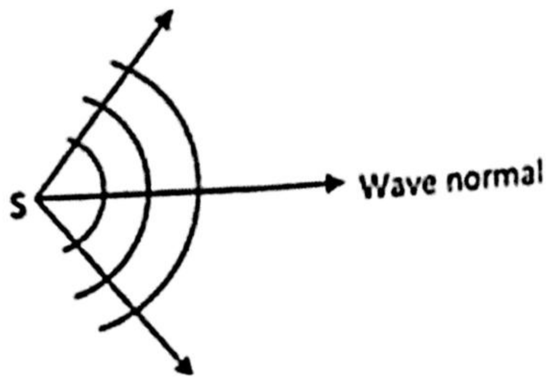


Figure : 8

4.9.4 Plane Wavefront

A wavefront in the form of plane surface is called plane wavefront. It is obtained by keeping point source at a focus of a convex lens or at a large distance from the point source.

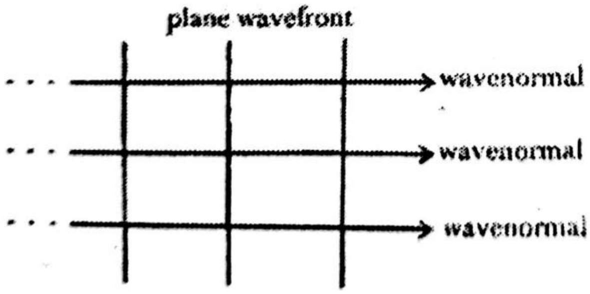


Figure : 9

4.9.5 Cylindrical Wavefront

If wavefront in the form of cylindrical surface than it is a cylindrical wavefront. It is obtained from an extended light source.

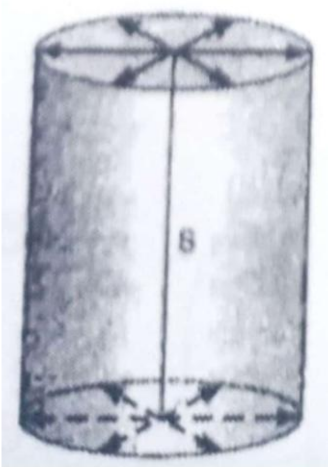


Figure : 10

4.9.6 Wave Normal

A normal drawn on the surface of the wavefront at any point in the direction of propagation of light is called a wave normal.

The ray of light shown in a plane wavefront or spherical wavefront is a wave normal. Wavefront transfers light energy in the direction perpendicular to its surface and is represented by a wave normal.

4.10 HUYGENS' PRINCIPLE

Huygens principle proposed by Christiaan Huygens in 1678 revolutionised our understanding of light and its characteristics. We are familiar with the rectilinear light theory that suggests light travels along straight paths. Huygens principle is one of the key methods for studying various optical phenomena.

4.10.1 Statement of Principle

Huygens' principle is stated as follows,

- Each point on the wavefront acts as a secondary source of light emitting secondary waves in all possible direction.
- The secondary waves progressing forward direction only taken to be effective.
- The locus or tangential surface to all these secondary waves at any instant gives new wavefront at that given instant.

4.10.2 Huygens' Construction of a Spherical Wavefront

(i) Let, PQ be a cross-section of a spherical wavefront emitted by a point source (S), at any instant. This can be called as primary wavefront.

(ii) Now consider points A, B, C on wavefront PQ. Thus, according to Huygens' Principle, they will act as secondary sources and emits secondary wavelets.

(iii) If 'c' is the speed of light in the medium, then in time 't', each wave will describe a distance 'ct' in forward direction as the secondary waves moving in the backward direction do not exist.

(iv) With points A, B, C as centres, circles can be drawn each of radius 'ct'. This each circle will represent a secondary wavefront.

(v) The common tangential surface P'Q' drawn to these secondary wavefronts represents the (new) position of the wavefront after time 't'.

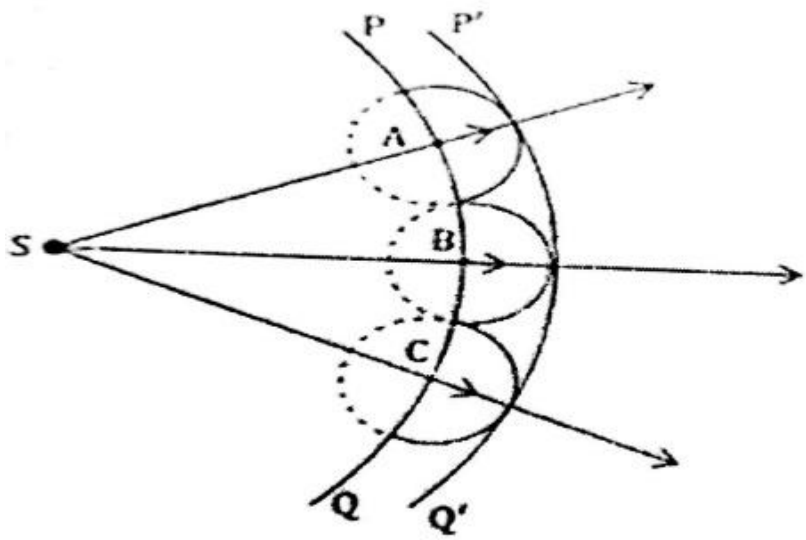


Figure : 11

4.10.3 Huygens' Construction of Plane Wavefront

(I) Let, PQ be a plane wavefront emitted by a point source (S) at any instant and at very large distance, this can be called as a primary wavefront.

(II) Now consider points A, B, C on wavefront PQ, They act as secondary sources, and send out secondary wavelets as per Huygens' principle.

(III) If 'c' is the speed of light in the medium in time 't', each wave will describe a distance 'ct' in forward direction as the secondary waves moving in the backward direction do not exist.

(IV) With A, B, C as centres, circles can be drawn each of radius 'ct', This each circle will represent a secondary wavefront.

(V) The common tangential surface P'Q', drawn to these secondary wavefronts represents the new position for the plane wavefront afater time 't'.

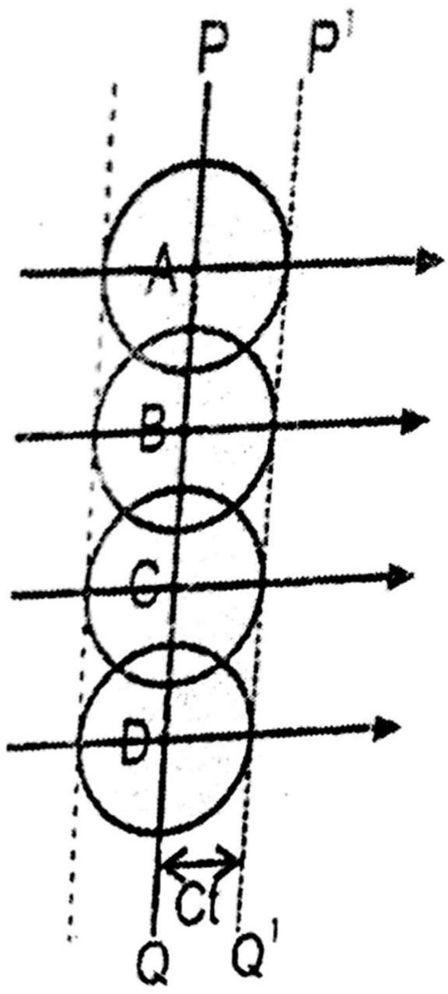


Figure : 12

SAQ :

Q12. State Huygens' principle and explain the Huygens' construction of a spherical wavefront.

Q13. Using Huygens' principle explain the construction of a plane wavefront.

Q14. In an isotropic medium, the secondary wavelets centred on every point of a given wavefront are all

(a) spherical

(b) cylindrical

(c) oval

(d) of arbitrary shape

Q15. A point source of light is kept at the focus of a convex lens. The wavefront emerging from the lens is :

(a) A plane wavefront

(b) A diverging wavefront

(c) A spherical wavefront

(d) A cylindrical wavefront

Q16. As a plane wavefront propagates its radius of curvature :

(a) decreases

(b) increases

(c) first increases and then decreases

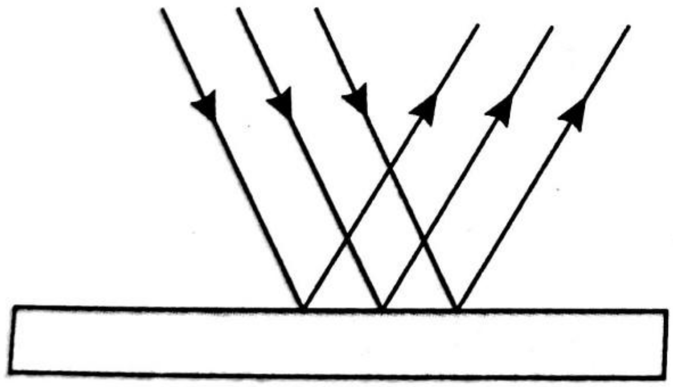
(d) becomes infinity

4.11 REFLECTION OF LIGHT

When light rays strike the boundary of two media such as air and glass, a part of light is turned back into the same medium. This is called reflection of light.

4.11.1 Regular Reflection

When the reflection takes place from a perfect plane surface it is called regular reflection [see Fig. 13 (a)]. In this case, the reflected light has large intensity in one direction and negligibly small intensity in other directions.

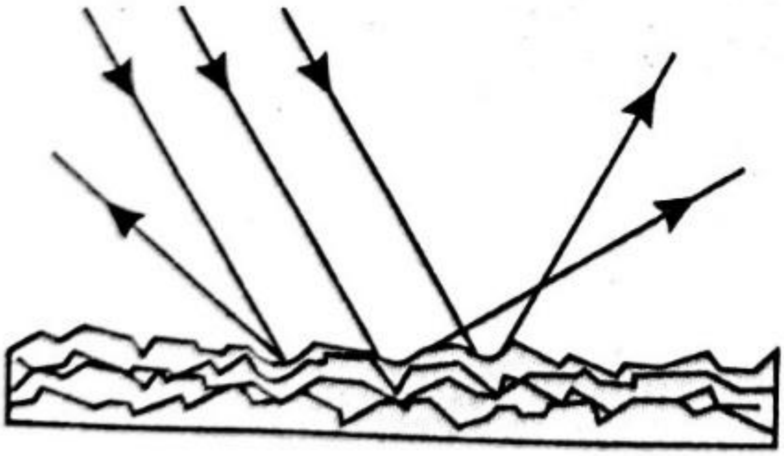


(a) Regular reflection

Figure : 13 (a)

4.11.2 Diffused Reflection

When the surface is rough, we do not get a regular behaviour of light. Although at each point light ray gets reflected irrespective of the overall nature of surface, difference is observed because even in a narrow beam of light there are many rays which are reflected from different points of surface. It is quite possible that these rays may move in different directions due to irregularity of the surface. This process enables us to see an object from any position. Such a reflection is called as diffused reflection [see Fig. 13 (b)]. For example, reflection from a wall, from a newspaper, etc. This is why you cannot see your face in the newspaper and on the wall.



(b) Diffused reflection

Figure : 13 (b)

4.11.3 Law of Reflection

Reflection of light is the process of deflecting a beam of light in the same medium. It has been found experimentally that rays undergoing reflection follow two laws called the Laws of reflection :

1. The incident ray, the reflected ray, and the normal at the point of incidence lie in the same plane. This plane is called the plane of incidence (or Plane of reflection).

2. The angle of incidence (the angle between normal and the incident ray) and the angle of reflection (the angle between the reflected ray and the normal) are equal, i.e. <i = <r.

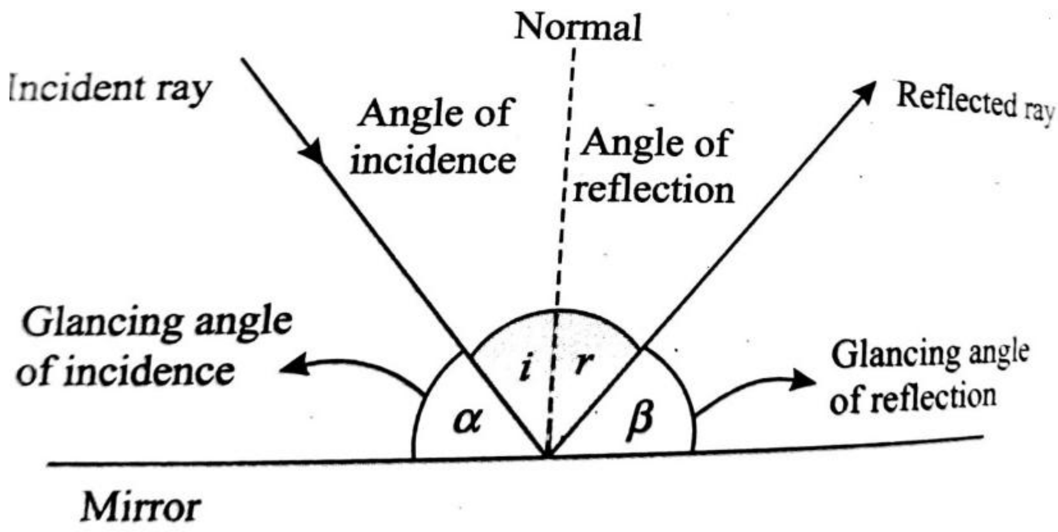
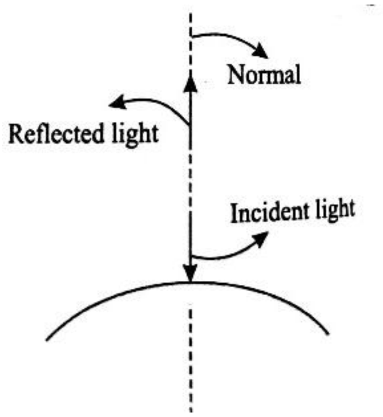
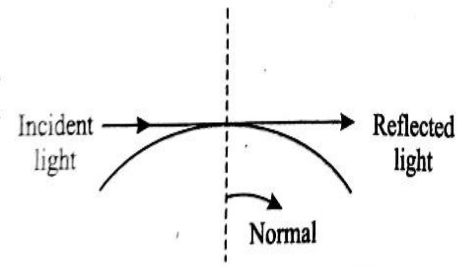


Figure : 14

Remember :

- The angle between an incident ray (on a plane of surface) and the surface is called the **glancing angle of incidence**. Simply it is angle made by the incident ray with the plane of the surface (on which the light is incident)
- The angle between a surface and reflected ray is called **glancing angle of reflection**. Simply it is the angle made by the reflected light with the plane of the surface (from which the light is reflected)
- In case of reflection,

  Glancing angle of incidence (α) = Glancing angle of reflection (β)

Special Cases	
Normal incidence : In case light is incident normally $i = r = 0$	
Grazing incidence : In case light strikes the reflecting surface tangentially $i = r = 90^o$	

4.12 REFRACTION OF LIGHT

Deviation or bending of light rays their original path while passing from one medium to another is called refraction. It is due to change in speed of light as light passes from one medium to another medium. If the light is incident normally then it goes to the second medium without bending, but still it is called refraction.

4.12.1 Refractive Index

Whenever light travels from vacuum to any other medium, its speed slow down. The speed of light is maximum in vacuum.

Why a light wave slows down when it enters an optical medium?

The answer can be found in electromagnetism where we find from Maxwell's equations that the speed of light in vacuum is given by,

$$C_{vac} = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$$

In a medium it is given by

$$C_{med} = \frac{1}{\sqrt{\varepsilon_r \varepsilon_0 \mu_r \mu_0}} = \frac{C_{vac}}{\sqrt{\varepsilon_r \mu_r}}$$

Since ϵ_r and μ_r are both greater than 1, speed of light in a medium will be less than the speed of light in vacuum. The ratio of the speed of light in vacuum to the speed of light in the medium is called the refractive index of the medium, represented by μ.

$$\mu = \frac{.c_{vac}}{C_{med}} = \frac{c}{v} = \frac{speed\ of\ light\ in\ vacuum}{speed\ of\ light\ in\ medium}$$

Refractive index of a medium is defined as the factor by which speed of light reduces as compared to the speed of light in vacuum. More (less) refractive index implies less (more) speed of light in that medium, which therefore is called denser (rarer) medium.

This μ (refractive index of medium w.r.t. vacuum) is called the absolute (or standard refractive index. Absolute refractive index is commonly referred to simply as refractive index. Further, refractive index is a measure of speed of light in a transparent medium or technically a measure of the optical density of the material. For example, the speed of light in water is less than that in air, so water is said to be optically denser. Optical density in general correlates with mass density.

However, in some cases, a material with greater optical density than another can have a lower mass density, e.g., mass density of turpentine oil is less than that of water but its optical density is higher. Thus, greater the refractive index of a material, the greater is the material's optical density and the smaller is the speed of light in that material (as $v = c/\mu$).

4.12.2 Laws of Refraction

The bending or the change in direction of propagation of light occurs except when it strikes the interface normally, i.e., when angle of incidence, i = 0.

The incident ray, as earlier said, is described by the angle of incidence whereas the refracted ray is describe by the angle of refraction (r), both angles measured from the normal to the interface of point of incidence (O).

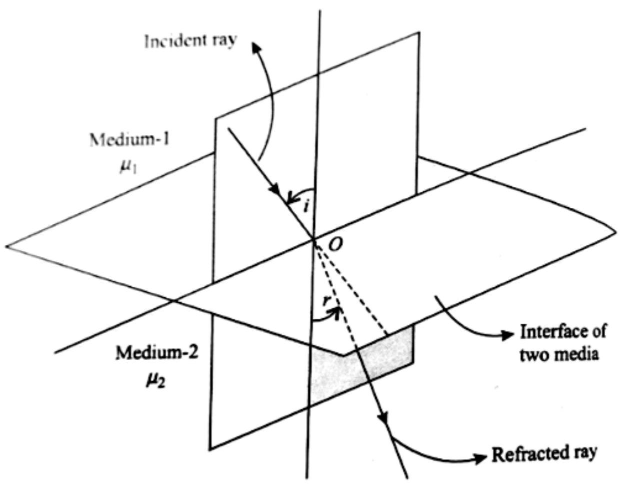


Figure : 15

Law 1 : The incident ray, the normal to the interface at the point of incidence and the refracted ray, lie in the same plane.

Law 2 : The ratio of the sine of the angle of incidence to the sine of the angle of refraction is always a constant (a different constant for different media),

For any two media the ratio of the sine of the angle of incidence to the sine of the angle of refraction is always constant.

$$\boxed{\frac{\sin i}{\sin r} = constant = \ {}^{1}\mu_2 = \frac{\mu_2}{\mu_1}} \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Equation (1) is known as Snell's law. The constant ${}^{1}\mu_2$ (or μ_{21}) is called the refractive index of medium-2 with respect to medium-1 when a rays of light travels from medium-1 to medium-2. In other words ${}^{1}\mu_2$ (or μ_{21}) is the refractive index of the medium in which the refracted ray lies w.r.t. the medium in which the incident ray lies.

Equation (1) can be written as $\mu_1 . \sin i = \mu_2 . \sin r = constant.$

The product of the refractive index and the sine of angle made by the ray with the normal at incident is constant for a given ray in both the media.

4.13. LAW OF REFLECTION ON THE BASIS OF HUYGENS' WAVE THEORY OF LIGHT

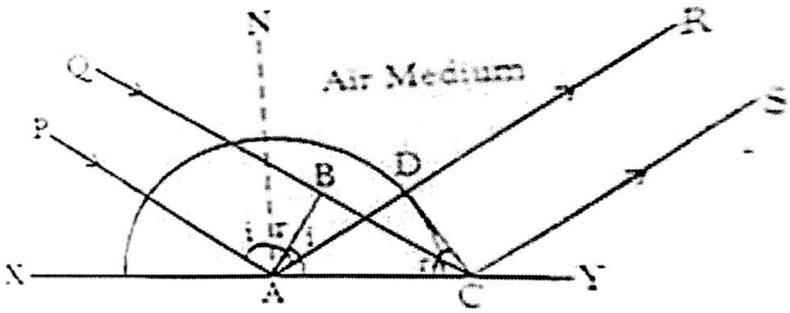


Figure : 16

From figure,

XY = Plane refracting surface,

NA = Normal drawn to XY,

PA and QC = incident light rays,

∠PAN = ∠BAC = ∠i = angle of incidence

AR and CS = reflected light rays,

∠RAN = ∠DCA = ∠r = angle of reflection

AB = incident plane wavefront,

CD = reflected plane wavefront.

Explanation of reflection of light from reflecting surface :

1. Let XY is plane reflecting surface.

2. Consider AB is a plane wavefront bounded by rays PA and QB. Let AB is incident obliquely on surface XY, in air medium.

3. Wavefront first reaches to point A and it act as secondary source of light and will emit secondary waves in air.

4. Suppose that the incident wavefront travels from B to C in time t.

   ∴ BC = c t

   Where c = Velocity of light in air.

5. During time t, secondary waves from point A, will covers equal distance to BC. So, Where, radius of secondary wavefront will be equal to BC.

6. Taking A as a centre and radius BC draw a hemisphere which represent secondary wavelets.

7. Draw a tangent CD to hemisphere.

8. The point C and D are in the same phase as light has travelled for equal time to reach this point. Hence CD represents the reflected wavefront bounded by rays AR and CS.

9. The hemisphere has a radius AD.

$$\therefore\ AD = BC = ct$$

$$From\ the\ Figure$$

$$\angle PAN = \angle BAC = \angle i$$

$$\angle RAN = DCA = \angle r$$

$$\boldsymbol{Also\ from\ figure}$$

$$\Delta\ ABC\ and\ \Delta\ ADC$$

$$\angle ABC = \angle ADC = 90^o$$

$$AC\ is\ common$$

$$BC = AD = ct$$

$$\therefore\ \Delta ABC \cong \Delta ADC$$

$$\therefore\ \angle BAC = \angle DCA$$

$$\angle i = \angle r$$

(a) Thus angle of incidence is equal to the angle of reflection.

(b) The incident ray, reflected ray and the normal lies in the same plane.

(c) The incident ray, the reflected ray lie on the opposite sides to that of normal.

Hence law of reflection is proved by Huygens' wave theory.

4.14 LAWS OF REFRACTION OF LIGHT FROM THE HUYGENS' WAVE THEORY OF LIGHT

From figure,

XY = Plane refracting surface,

NAM = Normal drawn to XY,

PA and QC = incident light ray,

∠PAN = ∠BAC = ∠i = angle of incidence

AR and CS = refracted light ay,

∠MAD = ∠DCA = ∠r = angle of refraction

AB = incident plane wavefront,

CD = refracted plane wavefront.

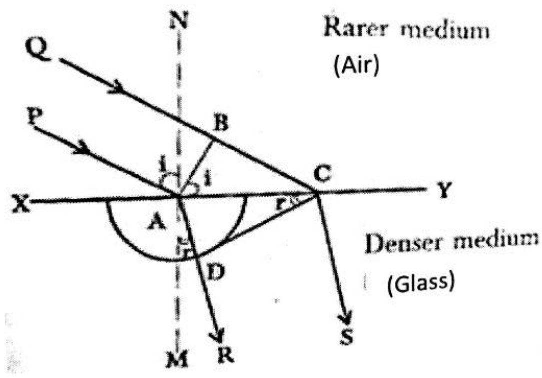


Figure : 17

Explanation of refraction of light from refracting surface :

1. Let XY be refracting surface separating rarer medium (air) and denser medium (glass).

2. Consider AB is a plane wavefront bounded by rays PA and QB. Let AB be incident obliquely on surface XY in air medium.

3. Wavefront first reaches to point A and acts as secondary source of light and will emit secondary waves in denser medium.

4. Suppose that the incident wavefront travels from B to C in time t.

$$\therefore BC = c_1 t$$

c_1 = Velocity of light in rarer medium.

5. During time t, secondary waves from point A will covers distance equal to c_2t in denser medium. Where, c_2 = velocity of light in denser medium.

6. Taking A as a centre and radius $c_2 t$, draw a hemisphere in the denser medium which represent secondary wavelets.

7. Draw a tangent CD to hemisphere.

8. The point C and D are in the same phase. Hence CD represents the refracted wavefront bounded by rays AR and CS.

9. The hemisphere has a radius AD.

$$\therefore AD = c_2 t$$

To Prove law of refraction :

From the fig.

$\angle i + \angle NAB = 90^o$ ………………………… (2)

$\angle NAB + \angle BAC = 90^o$ ………………………… (3)

From eq^n (2) and eq^n (3), we get

$\angle NAB + \angle BAC = \angle i + \angle NAB$

$\angle BAC = \angle i$ …………………………… (4)

Now, in ΔABC,

$\sin i = \frac{BC}{AC}$ …………………………… (5)

Also,

$\angle r + \angle DAC = 90^o$ …………………………… (6)

And in ΔADC,

$\angle DAC + \angle DCA = 90^o$ …………………………… (7)

From eq^n (5) and eq^n (6), We get

$\angle DAC = \angle r$

In ΔADC,

$\sin r = \frac{AD}{AC}$ …………………………… (8)

Dividing eq^n (5) and eq^n (8), we get

$$\frac{sin\ i}{\sin r} = \frac{BC}{AC} \times \frac{AC}{AD}$$

$$\frac{BC}{AD} = \frac{c_1 t}{c_2 t}$$

$$\therefore \frac{\sin i}{\sin r} = \frac{c_1}{c_2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

By definition of refractive indices,

$$I_{\mu_2} = \frac{\mu_2}{\mu_1} = \frac{C_1}{C_2} = \frac{velocity\ of\ light\ in\ rarer\ medium}{velocity\ of\ light\ in\ denser\ medium}$$

I_{μ_2} = Refractive index of Denser medium w.r.t rarer Medium.

Put in eq^n (9), We get

$$I_{\mu_2} = \frac{\mu_2}{\mu_1} = \frac{\sin i}{\sin r}$$

i.e.

$$\frac{\sin i}{\sin r} = \frac{C_1}{C_2} = \frac{\mu_2}{\mu_1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (10)$$

(i) Thus incident ray and refracted ray are on the opposite sides of the normal at the point of incidence and all three lie in same plane.

(ii) For given pair of media the ratio of the sine of the angle of incidence to sine of the angle of refraction is constant.

Hence law of refraction is proved by Huygens' wave theory.

To Show That $c_1 > c_2$:

From fig.

$\angle i > \angle r$;

sin i > sin r

$\frac{\sin i}{\sin r} > 1$

$\therefore\ \mu_2 > 1$

Refractive index of denser medium is always greater than 1.

$\therefore$ from equation (10) it is clear that, $c_1 > \ c_2$

Thus velocity of light in rarer medium is always greater than the velocity of light in denser medium.

SAQ:

Q17. Explain the phenomenon of reflection on the basis of Huygnes' wave theory of light.

OR

With a neat labelled diagram, explain reflection of light from a plane reflecting surface on the basis of wave theory of light.

Q18. Describe the laws of refraction of light from Huygens' wave theory of light.

OR

On the basis of Huygens' wave theory of light, prove that the velocity of light in a rarer medium is greater than the velocity of light in a denser medium.

OR

Explain refraction of light on the basis of wave theory. Hence prove the laws of refraction.

4.15. FERMATS PRINCIPLES

There was no simple relationship determined between the angles of refraction and incidence since 1621. Dutch investigator Willebrord Snell in 1621 claimed that it is the sines of the angles of refraction and incidence that uphold a constant ratio. Though the statement claimed was correct; however, the observation has not addressed any cause. However, Fermat was able to determine the cause of the refractive behavior of light. In 1662, Fermat demonstrated that a beam of light follows a path of the fastest time instead of a smaller distance. For the refraction and reflection at the plane surfaces, Fermat's Principle of Least Time holds. It means that a ray of light passing from one point to another chooses a certain path along which the time taken is minimum. However, for the spherical surfaces, the time taken by the light ray is either maximum or minimum.

4.15.1 Who Proposed the Fermat's Principle?

Fermat's principle was stated by Pierre de Fermat in 1662 and is suitable to study optical devices. Initially, the principle was a controversial statement as it appeared to assign knowledge of nature's intent. Until the 19th century, nobody was aware that alternative paths traveling from one point to another are a fundamental property of waves.

When two points A and B are specified, and a wave front is growing from point A, then it will sweep all the possible rays radiating from A. However, if the wave front reaches point B, then it will not only sweep all the ray paths with the same endpoints but also the infinitude of nearby paths.

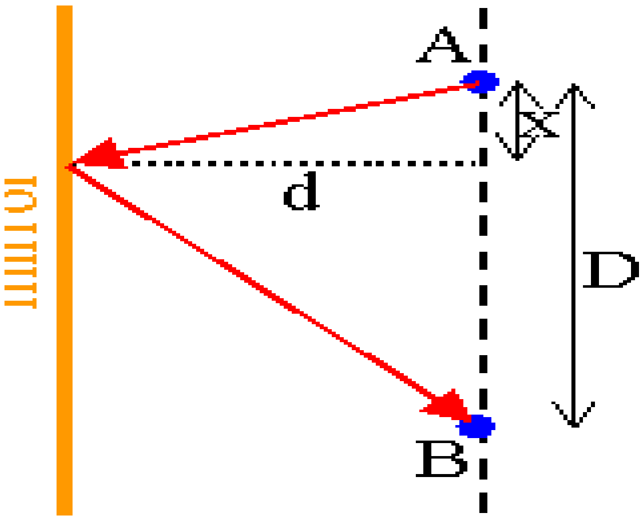


Figure : 18

The diagram shows that any light rays passing from one medium to another medium bend depending on the rarer and denser medium.

Fermat’s principle demonstrates that any ray that occurs from one point to another point will cover the smallest path possible.

State Fermat's Principle

According to Fermat’s principle, light traveling between two points pursues a path such that the optical length between the points is equal. The principle is the link between wave optics and ray optics.

One of the great ways to state Fermat's principle is that the path taken by light rays in traveling between two points involves either a maximum or minimum time. It means that two light rays deviating from a distant object will have the same optical path lengths. According to Fermat's principle of Least Time, the path traveled by light rays between the given two points is the actual path that can be traversed in the least time.

An Example of Fermat's Principle

Consider a beam of light traveling from point A to point B. Here, point A is in the air, and point B is in the glass. Fermat presented that the path of light rays is specified by the Principle of Least Time. Fermat's Principle of Least Time states that a light ray going from A to B will take a shorter time.

As the speed of light is identical everywhere along all the possible paths, the shortest path is the one that involves the shortest distance.

According to Fermat's principle, the entire path made by light to travel from A to B should satisfy the boundary conditions along with the condition that ray takes the path that requires the least amount of time.

Let the path taken by ray to travel from A to B is equal to ds. Here, ds refer to the small difference of the first order and it refers to the time difference taken along the path, that is;

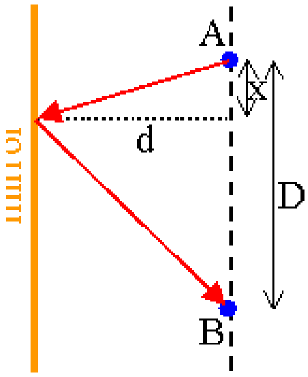


Figure : 19

$$\frac{dt}{ds} = 0$$

The diagram shows that the ray of light passing from point A that is in medium 1 to point B that is in medium 2 covers the shortest distance possible.

4.15.2 Application of Principle

There are several observations that can be made using Fermat’s principle. These principles will help us to prove and explore the realm of geometric optics.

- From Fermat's principle, it is found that in homogenous medium rays of light are rectilinear. It shows that light travels in a straight line in a medium having a constant index.

- The angle of incidence is equal to the angle of reflection, which is often called the Law of Reflection.

- Fermat’s principle refers to the fundamental law of optics that is used in the derivation of other laws of geometrical optics.
- One can make some useful observations about the conic surfaces based on Fermat’s principle.

For Example, two conjugate points are chosen that are perfect images of each other. The points are chosen in such a way that the optical path length of all paths connecting them is equal. In the case of an ellipse, the point source is located at one focus and is an image of the point located at another focus. However, in the case of the parabola, the point on one focus should be infinite.

The conic surfaces are very useful in mirror optics. The telescope is the best example of a conic surface and its usefulness in mirror optics. The telescope is designed using two conjugate points. Conjugate points are also known as focal points, these points can join one another roughly by 1 parameter. These points are perfect images of each other and all the rays connect equally at the optical path length of the conjugate points.

For Example, the north pole and the south pole are connected to each other on the sphere by any meridian.

Now let us consider an ellipse that has a conic surface. An ellipse is basically a plane curve that is surrounded by two different focal points. The sum of all the focal points on the curve will be constant.

Similarly, if we were to consider a parabola on a conic surface. The parabola is like the circle with a quadratic relation but unlike the circle, a parabola will not be squared at both x and y. Considering this we can interpret that an aggregate of rays will pass through the focus. The rays will be parallel to one another and also to the axis of the paraboloid after it has been recreated by the paraboloid.

For Example, The Newtonian Telescope is designed based on the same principle. It collects and focuses lights from distant objects. Hence a conic surface will always have two foci with optical conjugating points.

4.16 PERCEPTION OF LIGHT

These terms describe **the ability to perceive the difference between light and dark, or daylight and nighttime**. A person can have severely reduced vision and still be able to determine the difference between light and dark, or the general source and direction of a light.

4.17 SCATTERING OF LIGHT AND ITS IMPORTANCE

As sunlight travels through the earth's atmosphere, it gets scattered (changes its direction) by the atmospheric particles. Light of shorter wavelengths is scattered much more than light of longer wavelengths. (The amount of scattering is inversely proportional to the fourth power of the wavelength. This is known s Rayleigh scattering). Hence, the bluish colour predominates in a clear sky, since blue has a shorter wavelength than red and is scattered much more strongly. In fact, violet gets scattered even more than blue, having a shorter wavelength. But since our eyes are more sensitive to blue than violet, we see the sky blue.

Large particles like dust and water droplets present in the atmosphere behave differently. The relevant quantity here is the relative size of the wavelength of light λ, and the scatterer (of typical size, say, α), For a << λ, ibe gas Rayleigh scattering which is proportional to $1/\lambda^4$. For a >> λ, i.e. large scattering objects (for example, raindrops, large dust or ice particles) this is not true; all wavelengths are scattered nearly equally. Thus, clouds which have droplets of water with a >> λ are generally white.

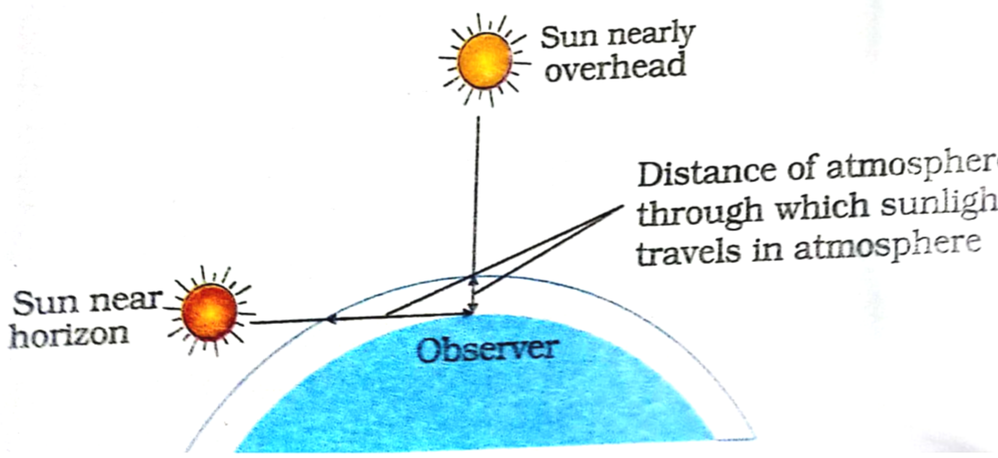


Figure : 20

At sunset or sunrise, the sun's rays have to pass through a larger distance in the atmosphere (Fig. 9.28). Most of the blue and other shorter wavelengths are removed by scattering. The least scattered light reaching our eyes, therefore, the sun looks reddish. This explains the reddish appearance of the sun and full moon near the horizon.

4.18 SUMMARY

1. Reflection is governed by the equation $\angle i = \angle r'$ and refraction by the Snell's law, sini/sinr = n, where the incident ray, reflected ray, refracted ray and normal lie in the same plane. Angles of incidence, reflection and refraction are *i, r'* and *r*, respectively.

2. The critical angle of incidence i_c for a ray incident from a denser to rarer medium, is that angle for which the angle of refraction is 90^o. For $i > i_c$, total internal reflection occurs. Multiple internal reflections in diamond ($i_c = 24.4^o$), totally reflecting prisms and mirage, are some examples of total internal reflection. Optical fibres consist of glass fibres coated with a thin layer of material of lower refractive index. Light incident at an angle at one end come out at the other, after multiple internal reflection, even if the fibre is bent.

4.19 TERMINAL QUESTIONS

1. What is a wavefront? How is it produced? Derive the lens formula for a thin lens on the basis of the wave theory of light.

2. Write short notes on :

   (a) Wave theory of light.

   (b) Huygens principle.

   (c) Newton's corpuscular theory.

3. Discuss the nature of light. How do you explain the phenomenon of reflection, refraction and rectilinear propagation of light on the basis of wave theory?

4. Write an essay on the nature of light.

5. What is Huygens principle? Obtain the laws of reflection and refraction on the basis of wave theory of light.

6. Deduce the laws of reflection with the help of Huygens theory of secondary wavelets.

7. What is Huygens principle? How would you explain the phenomenon of reflection and refraction of plane waves at surfaces on the basis of wave nature of light?

8. State and explain Huygens principle of secondary waves.

4.20 SOLUTION & ANSWERS OF TERMINAL QUESTIONS

1. Section 4.9.2

2. (a) Section 4.8

   (b) Section 4.10

   (c) Section 4.5

3. Section 4.4, 4.11, 4.12

4. Section 4.3, 4.4

5. Section 4.10, 4.13, 4.14

6. Section 4.13

7. Section 4.10, 4.13, 4.14

8. Section 4.10

4.21 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.

2. Optics – Ajoy Ghatak

3. Optics – Principles and Application, K. K. Sharma, Academic Press.

4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.

5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT- 5 CONCEPT OF POLARIZATION

STRUCTURE:

5.1 INTRODUCTION:

- Interference and diffraction phenomena proved that light is a wave motion. These phenomena are used to find wavelength of light. However, they do not give any indication regarding the character of waves.

- Maxwell developed electromagnetic theory and suggested that light-waves are electromagnetic waves. Electromagnetic waves are transverse waves, so it is obvious that light waves are also transverse waves.

- Longitudinal waves are waves in which particles of medium oscillate along the direction of propagation of wave (e.g., sound wave).

- Transverse waves are waves in which particles of medium oscillate perpendicular to the direction of propagation of wave. (e.g. Electromagnetic waves.)

- Polarization is possible in transverse wave.

- Unpolarized Light is the light is which the planes of vibration are symmetrically distributed about the propagation direction of the wave.

- Plane Polarized light is a wave in which the electric vector is everywhere confined to a single plane.

- A linearly polarized light wave is a wave in which the electric vector oscillates in a given constant orientation.

5.2 OBJECTIVE:

After studying this unit, student should be able to -

- Understand the concept of Polarization.
- Define the term plane of Vibration and Plane of Polarization.
- Explain the concept of Brewster law and Law of Malus.
- Understand the Hygens theory of double reflection.
- Define Polarizer and Analyzer.

5.3 WHAT IS POLARIZATION?

Polarization, in Physics, is defined as a phenomenon caused due to the wave nature of electromagnetic radiation. Sunlight travels through the vacuum to reach the Earth, which is an example of an electromagnetic wave. These waves are called electromagnetic waves because they form when an electric field that interacts with a magnetic field. In this article, you will learn about two types of waves, transverse waves, and longitudinal waves. You will also learn about polarization and plane polarised light.

5.3.1 Transverse Waves

Transverse waves are waves, i.e., movement of the particles in the wave is perpendicular to the direction of motion of the wave.

Example 1: ripples in water, when you throw a stone.

Example 2: the motion of sound waves through the air.

5.3.2 Longitudinal Waves

Longitudinal waves are when the particles of the medium travel in the direction of motion of the waves.

Light is the interaction of electric and magnetic fields travelling through space. The electric and magnetic vibrations of a light wave occur perpendicularly to each other. The electric field moves in one direction and magnetic in another though always perpendicularly. So, we have one plane occupied by an electric field, the magnetic field perpendicular to it, and the direction of travel which is perpendicular to both. These electric and magnetic vibrations can occur in numerous planes. A light wave that is vibrating in more than one plane is known as unpolarized light. The light emitted by the sun, by a lamp or a tube light are all unpolarised light sources. As you can see in the image below, the direction of propagation is constant, but the planes on which the amplitude occurs is changing.

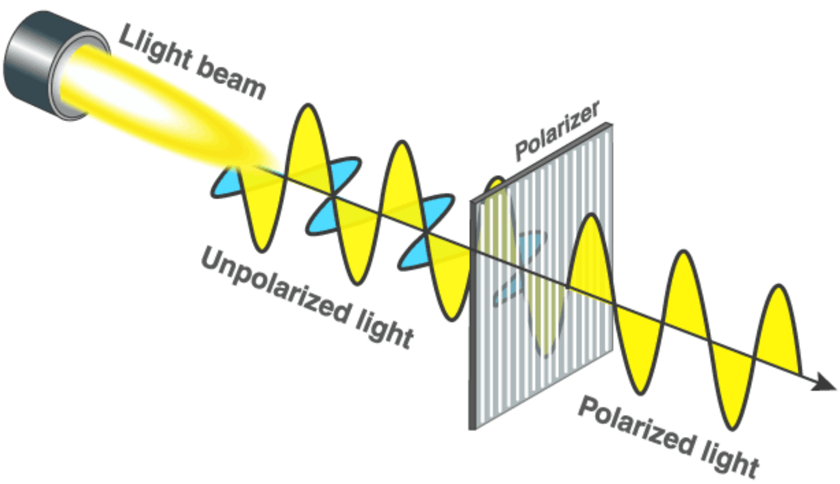


Figure : 1

The other kind of wave is a polarized wave. Polarized waves are light waves in which the vibrations occur in a single plane. Plane polarized light consists of waves in which the direction of vibration is the same for all waves. In the image above, you can see that a Plane polarized light vibrates on only one plane. The process of transforming unpolarized light into the polarized light is known as polarization. The devices like the purple blocks you see are used for the polarization of light.

5.4 SOME USEFUL TERMS AND DEFINITIONS

There are same important terms as follows :

5.4.1 Polarization of light

It is the phenomenon, due to which the vibration of light are restricted in a particular plane, is called the polarization of light.

When ordinary light is passed through a tourmaline crystal, it gets polarized, i.e., its vibrations originally present in all directions perpendicular to the direction of propagation are now confined to only one of these directions which are parallel to its crystallographic axis AB. This plane contains the direction of vibration and the direction of propagation of light.

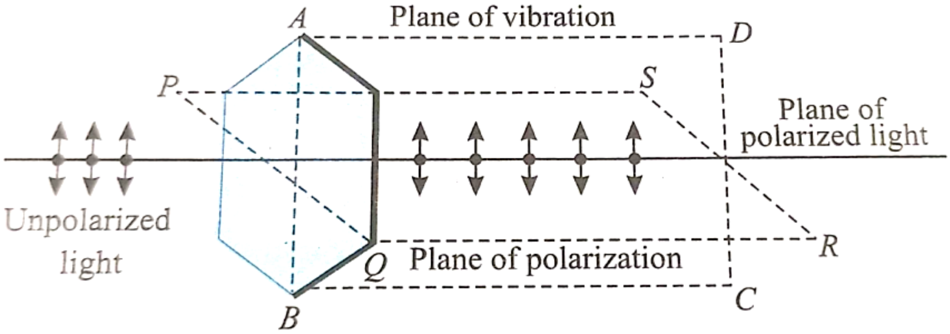


Figure : 2

5.4.2 Plane of Vibration

The plane within which the vibration of the polarized light are confined is known as the **plane of vibration**. It is represented by the plane (ABCD) in figure.

5.4.3 Plane of Polarization

It is the plane at right angles to the plane of vibration and passing through the direction of propagation of light is known as the **plane of polarization**. The plane (PQRS) perpendicular to the plane of vibration is called the plane of polarization.

5.4.4 Plane Polarized Light

It may be defined as the light, in which the vibration of the light (vibrations of the electric vector) is restricted to a particular plane.

According to the electromagnetic theory of light, the electric vector acts as light vector. Therefore, in a plane polarized light, the electric vector vibrates along a fixed line in a plane perpendicular to the direction of propagation.

In Figure 3 (a), the plane polarized component whose vibrations are confined to the plane of paper but perpendicular to the direction of propagation, has been represented. The plane of polarization is therefore perpendicular to the plane of the paper.

Figure 3 (b), on the other hand, is the representation of the other plane polarized component whose plane of polarization is in the plane of the paper and the vibrations which are perpendicular to the plane of the paper are represented by the dots on the direction of propagation.

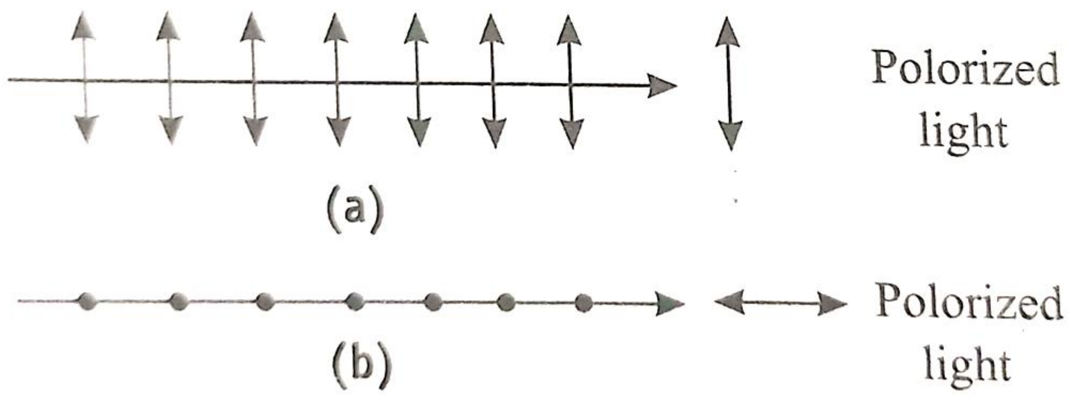


Figure : 3 (a), (b)

5.5 DETECTING PLANE POLARIZED LIGHT

The human eye lacks the ability to distinguish between randomly oriented and polarized light, and plane-polarized light can only be detected **through an intensity or color effect**, for example, by reduced glare when wearing polarized sun glasses.

5.5.1 Analyzer

We cannot make distinction between the unpolarized light and the plane polarized light through the naked eye or the polarizer alone. Another such crystal required to analyze the nature of light is called **analyzer**.

Analyzer is a device, which is used to find whether the light is polarized or unpolarized.

Both polarizer and analyzer are fabricated in the same way and wave the same affect on the incident light.

5.5.2 Polarizer

A tourmaline crystal or a Nicol prism used to produce plane polarized light is called **polarizer**. If the polarizer is rotated in the path of the ordinary light, the intensity of the light transmitted from the polarizer remains unchanged. It is because, in each orientation of the polarizer, the plane polarized light is obtained, which has vibrations in a direction parallel to the axis of the crystal in that orientation.

It is an optical device that transforms unpolarized light into polarized light. If it produces linearly polarized light. It is called a lineally Polarizer.

If natural light is incident on a linear polarizer, only that vibration which is parallel to the transmission axis is allowed to pass through the polarizer while the vibration that is in a perpendicular direction is totally blocked.

If we rotate the analyzer in the path of the light transmitted from the polarizer, so that the axes of the polarizer and the analyzer are parallel to each other, then the intensity of lights is found to remain unaffected (or maximum) as shown in Fig. 4 (a).

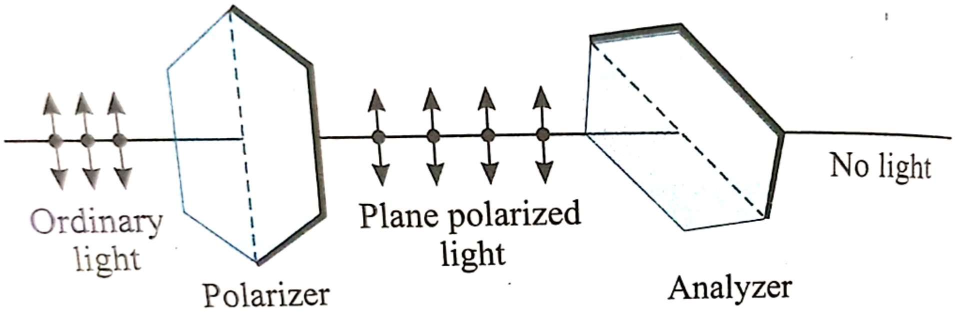


Figure : 4 (a)

But when the axis of the polarizer and the analyzer are perpendicular to each other, the intensity of light becomes minimum. In the positions shown in Fig. (b), the polarizer and the analyzer are said to be in **crossed positions**.

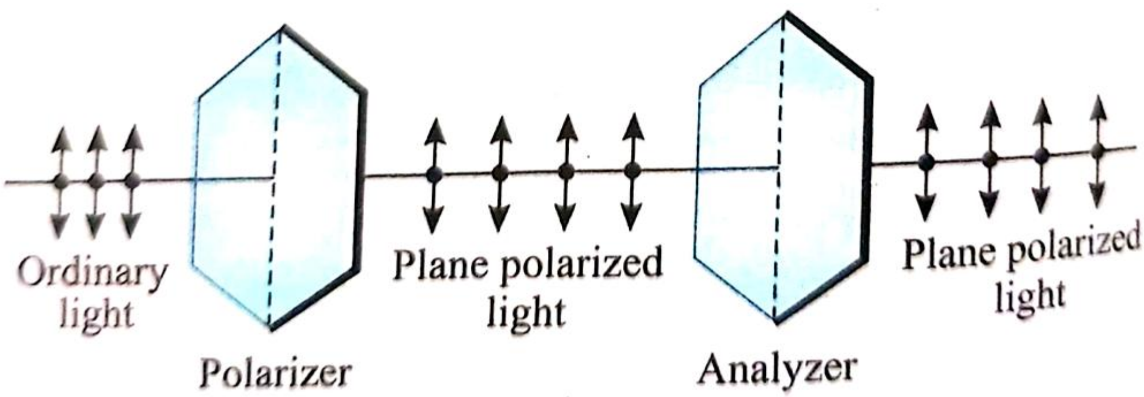
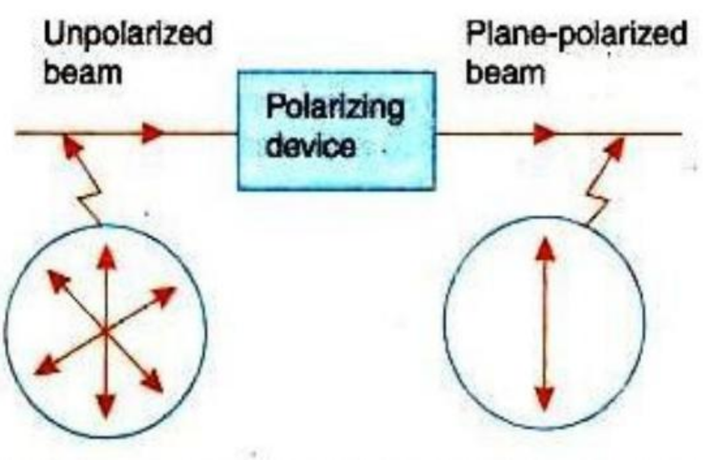


Figure : 4 (b)



A polarizer transforms unpolarized light into polarized light.

5.6 PRODUCTION OF LINEARLY POLARIZED LIGHT

Linearly polarized light may be produced from unpolarized light using following optical phenomena.

(i) Reflection

(ii) Refraction

(iii) Scattering

(iv) Selective absorption

(v) Double reflection.

Polarized light has many important applications in industry and engineering. One of the most important applications is in liquid crystal displays (LCDs) which are widely used in wristwatches, calculators, T.V. Screens etc. An understanding of polarization is essential for understanding the propagation of electromagnetic waves guided through wave-guides and optical fibers.

5.6.1 POLARIZATION BY REFLECTION

In 1808, E. L. Malus discovered the polarization of natural light by reflection from the surface of glass. He noticed that when natural light is incident on a plane sheet of glass at a certain angle, the reflected beam is plane polarized.

Figure 1 shows an unpolarized light beam AB incident on a glass surface. The incident ray AB and the normal NBN' represent the plane of incidence. The incident ray AB resolved into two components. (i) The perpendicular component, and(ii) parallel component. The perpendicular component is represented by dots and called the s-component. The parallel component is represented by the arrows and is called the p-component.

The reflected ray BC has predominance of s-component and it is partially polarized as show in fig(a). As the angle of the incident is varied, the polarization is also varied. At a particular angle θ_P, the reflection coefficient for p-component goes to zero and the reflected beam does not contain any p-component. It contains only s- component and is totally *linearly polarized*. The angle θ_P is called the polarizing angle.

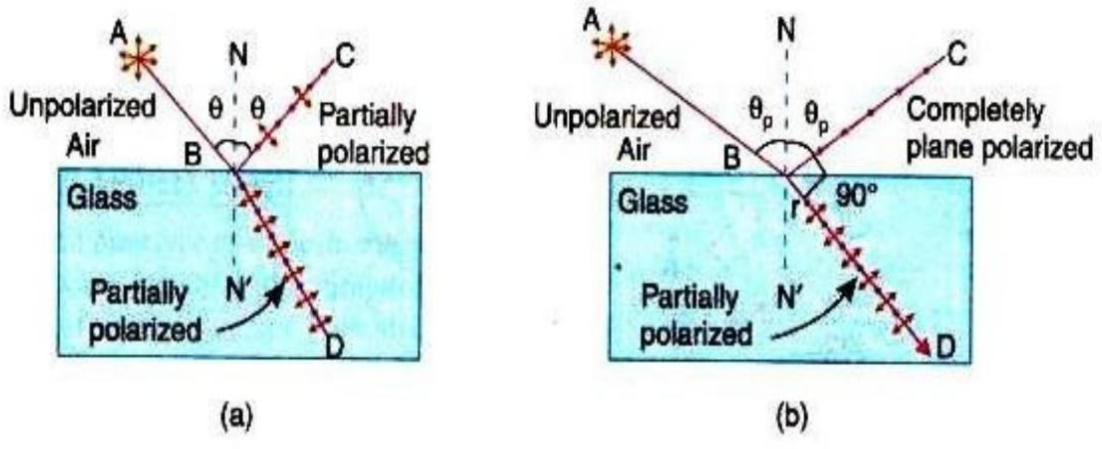


Figure : 5 (a) & (b)

In the path of 'BC', place a tourmaline crystal and rotate it slowly. It will be observed that light is completely extinguished only at one particular angle of incidence. This angle of incidence is known as the polarizing angle.

5.6.2 Brewster's Law

In 1892, Brewster performed number of experiments to study the polarization of light by reflection at the surfaces of different media.

He found that ordinary light is completely polarized in the plane of incidence when it gets reflected from a transparent medium at a particular angle known as the *'angle of polarization.'*

He proved that *'the tangent of the angle of polarization is numerically equal to the refractive index of the medium.'* Also, the reflected and retracted rays are perpendicular to each other.

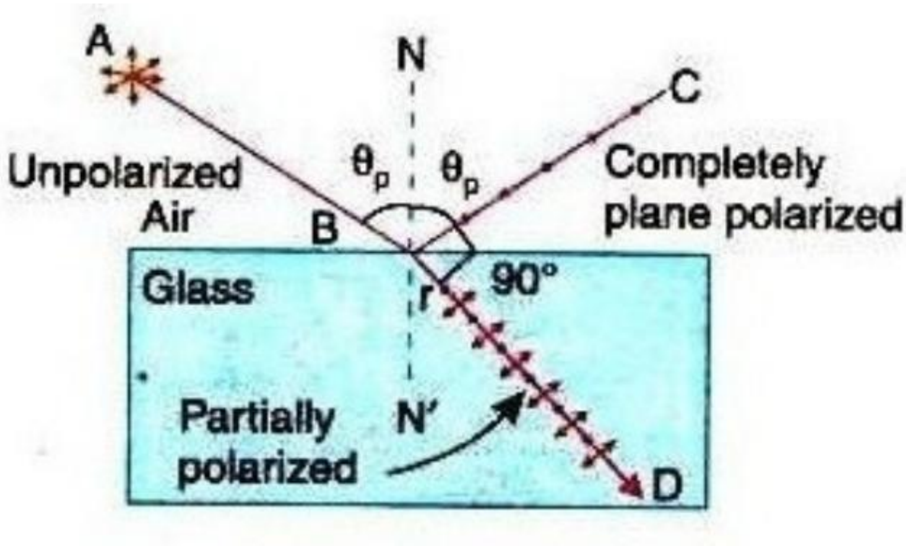


Figure : 6

If θ_P is the angle and μ is the refractive index of the medium, then

$$\mu = \tan\theta_p$$

This is known as Brewster's law

If natural light is incident on smooth surface at the polarizing angle, it is reflected along BC and refracted along BD as shown in figure 6. Brewster found that the maximum polarization of reflected ray occurs when it is at right angles to the refracted ray. It means that $\theta_p + r = 90^o$

According to Snell's law,

$$\frac{sin\theta_p}{sinr} = \frac{\mu_2}{\mu_1}$$

Where, μ_2 is the absolute refractive index of reflecting surface and μ_1

is the refractive index of the surrounding medium.

Hence,

$$\frac{sin\theta_p}{sin(90^0 - \theta_p)} = \frac{\mu_2}{\mu_1}$$

$$\frac{sin\theta_p}{cos\theta_p} = \frac{\mu_2}{\mu_1}$$

$$tan\theta_p = \frac{\mu_2}{\mu_1}$$

Above equation shows that the polarizing angle depends on the refractive index of the reflecting surface. The polarizing angle θ_p is also known as Brewster angle and denoted by θ_B. The light reflected from any angle other than Brewster angle is partially polarized.

As shown in the figure, unpolarized light is incident at our angle equal to the polarizing angle on the glass surface. It is reflected along 'BC' and retracted along 'BD'.

From Snell’s law,

$\mu = \frac{\sin i}{\sin r}$

(1)

From Brewster's law,

$\mu = \tan i = \frac{\sin i}{cons\ i}$

(2)

Comparing (1) & (2),

$$\cos i = \sin r = cos\left(\frac{\pi}{2} - r\right)$$

$$\therefore i = \frac{\pi}{2} - r$$

$$\therefore i + r = \frac{\pi}{2}$$

As $i + r = \frac{\pi}{2}$ angle CBD is also equal to $\frac{\pi}{2}$.

Therefore, the reflected and the retracted rays are at right angles to each other.

From Brewster's law, the value of 'i' for crown glass of refractive index 1.52 is given by.

$$i = tan^{-1}(1.52)$$

$$\therefore i = 56.7^{o}$$

For ordinary glass the approximate value for the polarizing angle is 57^{0}. For a refractive index of 1.7, the polarizing angle is 59.5^{0}. Thus, the polarizing angle is not widely different for different glasses.

As the refractive index of a substance varies with the wavelength of the incident light, the polarizing angle will be different for light of different wave length.

5.6.3 Applications of Brewster's Law

1. Brewster's law can be used to determine the reflective Indies of opaque materials.

2. It is used to calculate the polarizing angle for total polarization of reflected light, if reflective index of the material is known.

3. Brewster's angle can be utilized for transmitting a light beam in into or out of an optical fiber without reflections losses.

5.7 POLARIZATION BY REFRACTION

we know that light bends when it moves along different mediums, this property of light is called refraction. Hence we can conclude that **the polarization of light due to a pile of plates is called the polarization by refraction**.

5.7.1 Pile of Plates

When unpolarized light is incident at Brewster's angle on a smooth glass surface, the reflected light is totally polarized while the refracted light is partially polarized. If natural light is transmitted through a single plate, they it is partially polarized. If a stack of glass plates is used instead of a single plate is used instead of a single plate, reflections from successive surfaces of each glass plate filter the perpendicular component from the transmitted ray. So, transmitted ray consists of only parallel components.

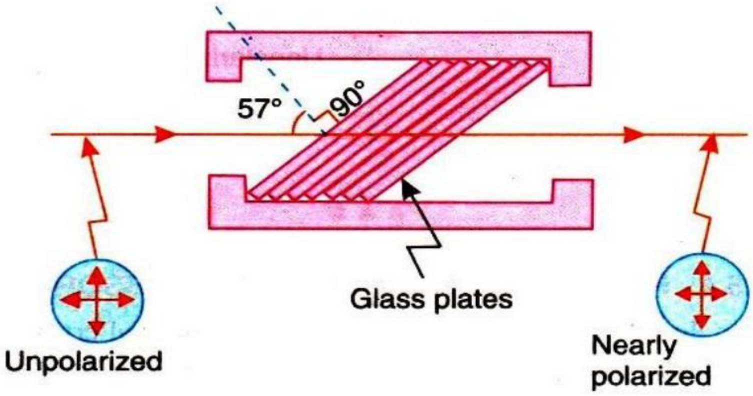


Figure : 7

I_p- Intensity of parallel component of refracted light.

I_s- intensity of perpendicular component of light.

Thus, degree of polarization of refracted (transmitted) light is given by

$$P = \frac{I_p - I_s}{I_p - I_s} = \frac{m}{m + \left(\frac{2\mu}{1-\mu^2}\right)^2}$$

Where m = no. of plates required.

μ= refractive index of material.

About 15 glass plates are required for this purpose. Thus, glass plates are kept inclined at an angle of 33^0 to the axis of the tube as shown in figure 7. Such an arrangement is called a pile of plates. When unpolarized right is incident on the plates at Brewster angle, the transmitted light will be polarized and parallel to the plane of incidence.

The drawback of this method is good portion of light is lost in reflections.

5.8 POLARIZATION BY SCATTERING

If a narrow beam of natural light is incident on a transparent medium, contain a suspension of ultramicroscopic particles, the scattered light is partially polarized. The degree of polarization depends on the angle of scattering. The beam scattered at 90^0 with respect to the incident direction is linearly polarized.

The direction of vibration of E vector in the scattered light will be perpendicular to the plane of propagation of light. Sun light scattered by air molecules is polarized. The maximum effect is observed on a clear day when the sun is near the horizon. The light reaching on the ground from directly overhead is polarized to the extent of 70% to 80%.

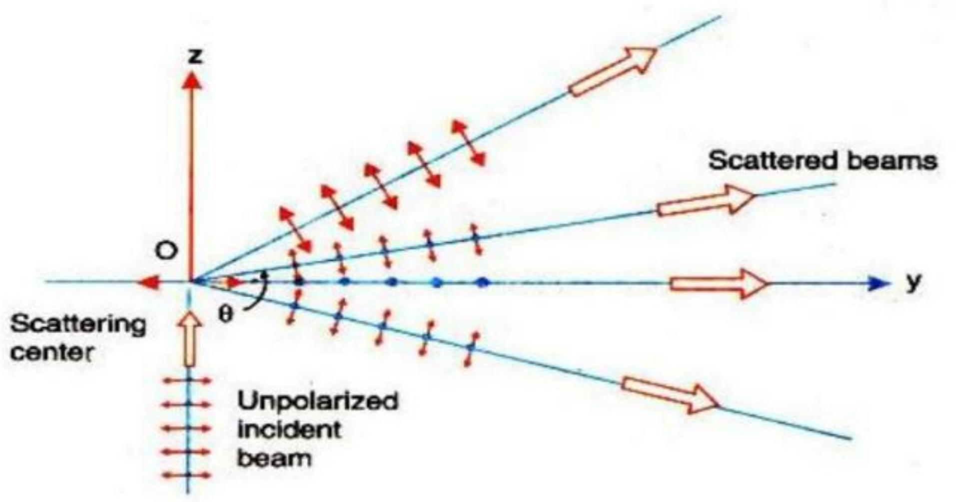


Figure : 8

5.9 POLARIZATION BY SELECTIVE ABSORPTION

In 1815 Biot discovered that certain mineral crystal absorbs light selectively. When natural light passes through a crystal such as tourmaline, it splits into two components which are polarized in mutually perpendicular places.

The crystal strongly absorbs light that is polarized in a direction parallel to a particular plane in the crystal but freely transmits the light component polarized in a perpendicular direction. This difference in the absorption for the rays is known as selective absorption or dichroism.

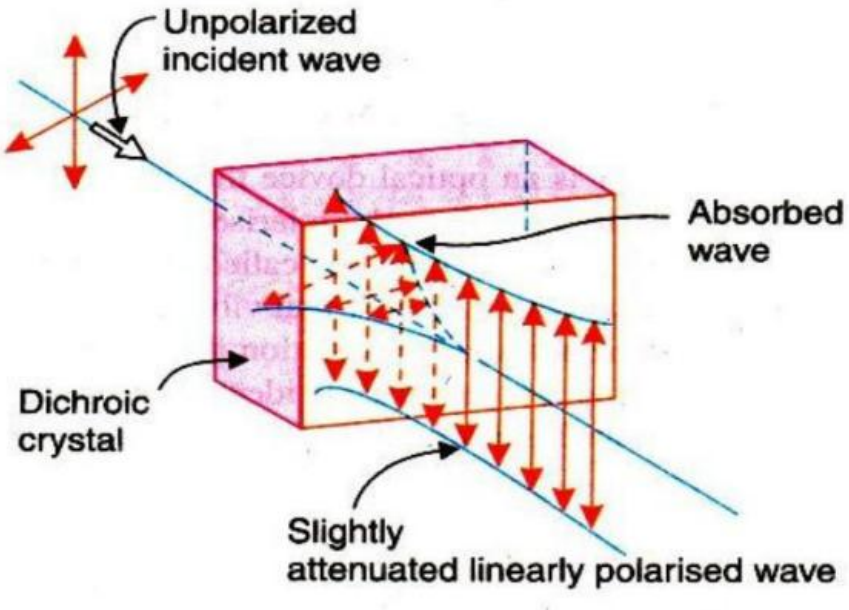


Figure : 9

If the crystal is of proper thickness, one of the components is totally absorbed and the other component emerging from the crystal is linearly polarized. It is shown in the figure 5.

The difference in absorption in different direction may be understood from the electron theory. When the frequency of incident light wave is close to natural frequency of the electron cloud, the light waves are absorbed strongly. Crystals that exhibit selective absorption are anisotropic.

The crystal splits the incident wave in to two waves. The component having its vibration perpendicular to the principal plane of the crystal gets absorbed. The component with parallel vibrations is less absorbed and it is transmitted. The transmitted light is linearly polarized. The drawback of this method is that the crystal of bigger size cannot be grown.

5.10 POLARIZATION BY DOUBLE REFRACTION

This phenomenon was discovered by Erasmus Bartholinus in 1969. When light is incident on a calcite crystal, it splits into two refracted rays. This phenomenon is called double refraction or birefringence. The crystal is called birefringent.

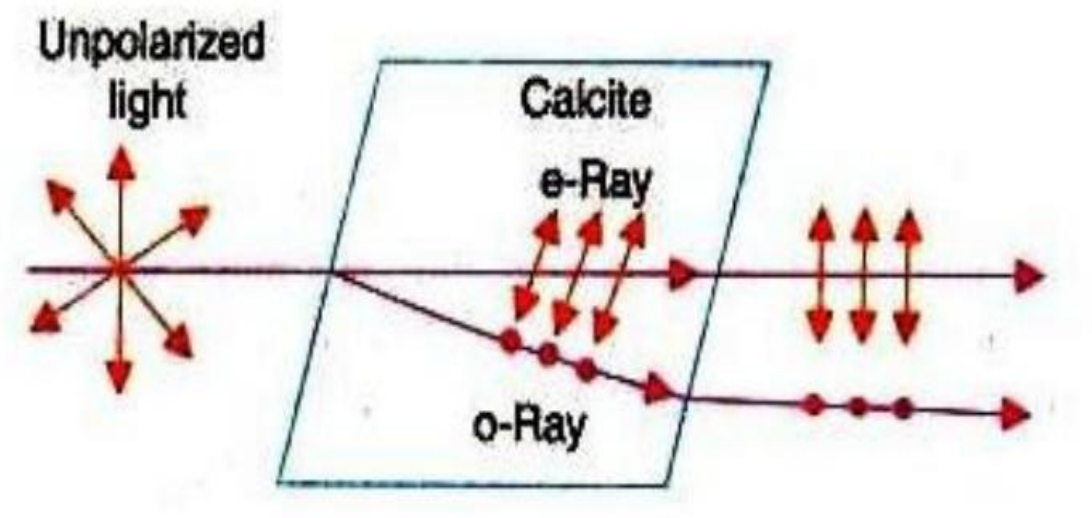


Figure : 10

The two rays produced in double refraction are linearly polarized in mutually perpendicular directions. The ray which obeys Snell's law of refraction is known as ordinary ray or o-ray. The other ray does not obey Snell's law is called extraordinary ray or e-ray.

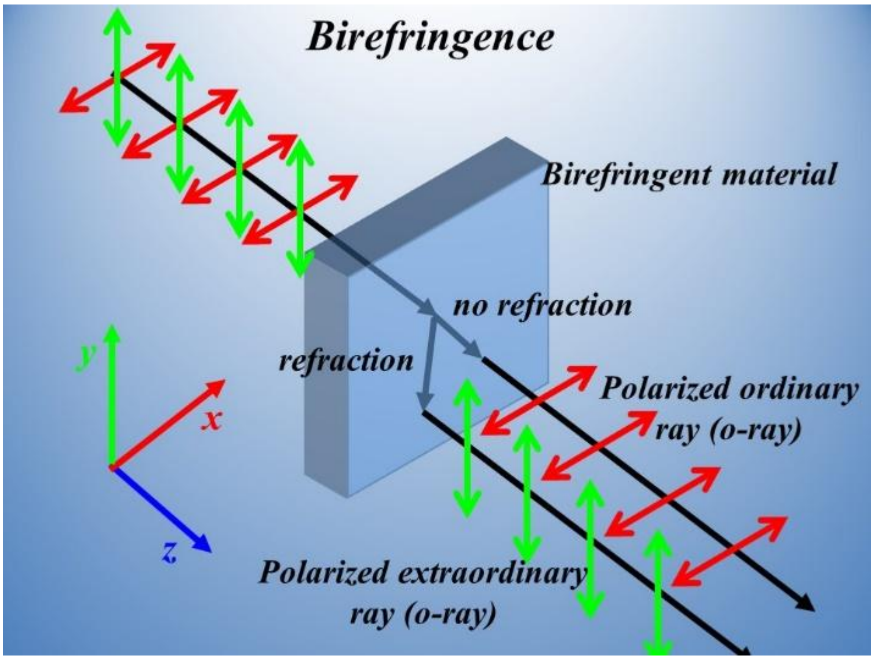


Figure : 11

5.11 NICOL'S PRISM

A Nicol prism is made from calcite crystal. It was designed by William Nicol in 1820.

A calcite crystal whose length is three times as its width is taken. The end faces of this crystals are ground3d in such a way that the angles in the principal section becomes 68º and 112º instead of 71º and 109º the crystal is cut in two pieces by a plane perpendicular to the principal section as well as the new end faces. The two parts of the crystal are then cemented together with Canada balsam. The refractive index of Canada balsam lies between the refractive indices for the ordinary and extra-ordinary rays for calcite.

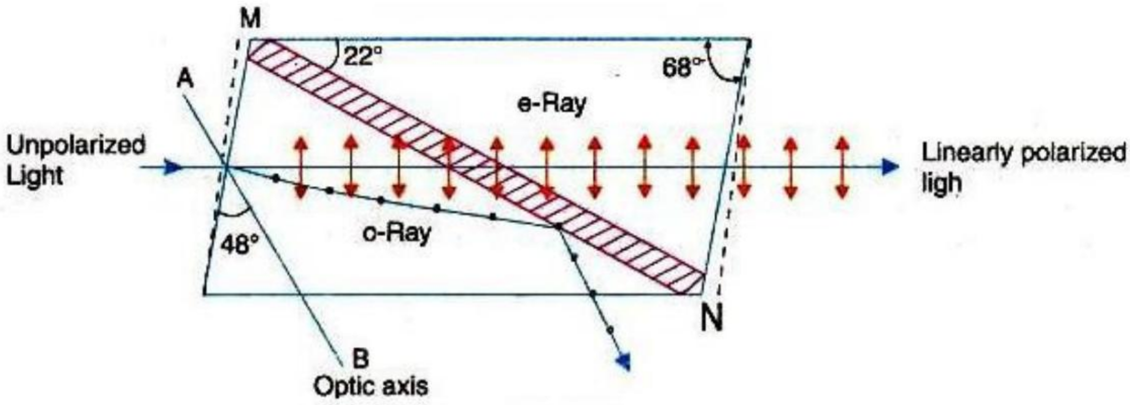


Figure : 12

($\mu_0 = 1.66$, $\mu_e = 1.486$ and $\mu_{Canada\ balms} = 1.55$)

The position of optic axis AB is as shown in figure.

Unpolarized light is made to fall on the crystal as shown in figure. The ray after entering the crystal suffers double refraction and splits up into O-ray and e-ray. The values of the refractive indices and the angles of incidence at the Canada balsam layer are such that are e-ray is transmitted while the o-ray is internally reflected. The face where the o-ray is incident-is black so that the o-ray is completely absorbed. Thus we get only the linearly polarized e-ray coming out of the Nicol prism with the direction of vibration as shown in figure. Thus, the Nicol Prism works as a polarizer.

The Nicol prism is the most widely used polarizing device. It is good polarizer but expressive. It has a limited field of view of about 28^o.

5.11.1 Polaroid Sheets

In 1928 E.H. Laud invented a method of aligning small crystal to obtain large polarizing sheets. The sheets are called Polaroid sheets.

Constructions:

- A clear plastic sheet of long chain molecules of polyvinyl alcohol (PVA) isheated.
- It is then stretched in a given direction to many times its original length.
- During this process, PVA molecules become aligned along the direction ofstretching.
- The sheet is there laminated to a rigid sheet of plastic.
- It is then exposed to iodine vapour.
- The iodine atoms attach themselves to the straight long chain of PVAmolecules.
- The conduction electrons associated with iodine can move along the chains.
- A sheet fabricated according to this process is known as H-sheet.

5.11.2 Working of Nicol Prism as Polarizer and Analyzer

Nicol Prism was invented by William Nicol in 1828. It is an optical device used for producing and analyzing plane polarized light. When a beam of light passes through a calcite crystal, it breaks up into o-ray and E-ray.

The Nicol prism is designed in such a way that it eliminates O-ray. Hence E- ray are transmitted through the prism.

It is used for the production and detection of plane polarized light.

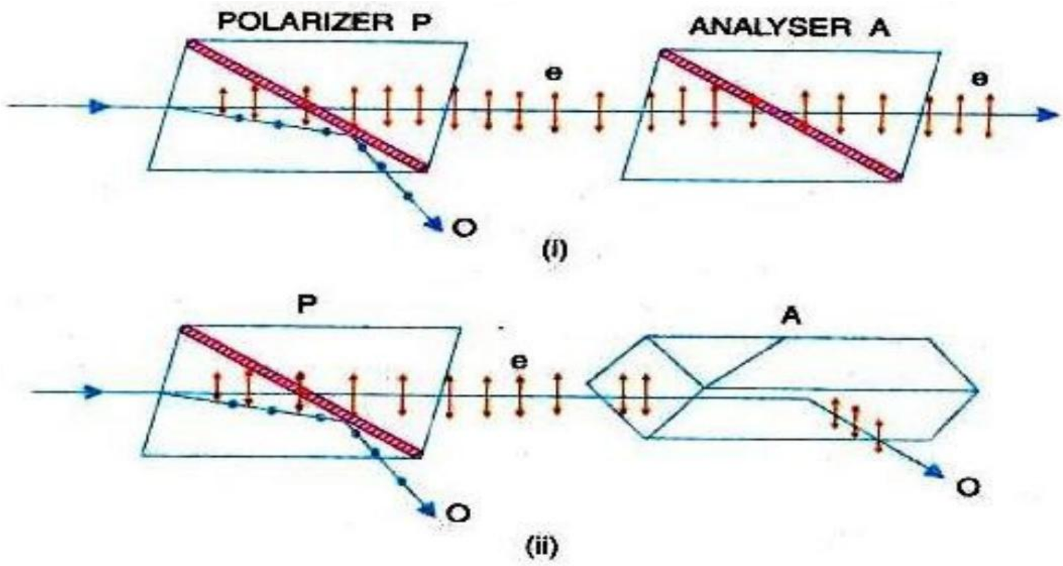


Figure : 13 (i) & (ii)

In above figure two Nicol Prisms are shown placed nearby to each other 'P', works as polarizer and it produces plane polarized light 'A' works as an analyzer and it analyses whether the emerging rays from 'P', is plane polarized, circularly polarized or elliptically polarized.

As shown in fig-13(i) o-ray is reflected by 'P1' would E-ray is transmitted through it. As shown in fig.-13 (ii) if 'P2' is gradually rotated the intensity of E-ray decreases and when the two prisms are crossed, then no light comes out of 'P2'.

It means that light coming out of 'P1' is plane polarized. When the polarized E-ray enters 'P2' it acts as an o-ray and it is then totally internally reflected by the Canada balsam layer. Hence light will not come out of 'P1'. Therefore, the prism 'P1' produces plane polarized light and prism 'P2' detect (analyze) it.

5.11.3 Polaroid Sheets as Polarizer and Analyzer

Fig. shows the Polaroid sheets in three different configurations.

- Fig 14(a) - transmission axis of the analyzer. A is parallel to polarizer P so light passes through the analyzer.
- Fig 14(b) Transmission axis is at angle O so light partially transmitted.
- Fig 14 (c) When the axes are perpendicular to each other, no light is transmitted.

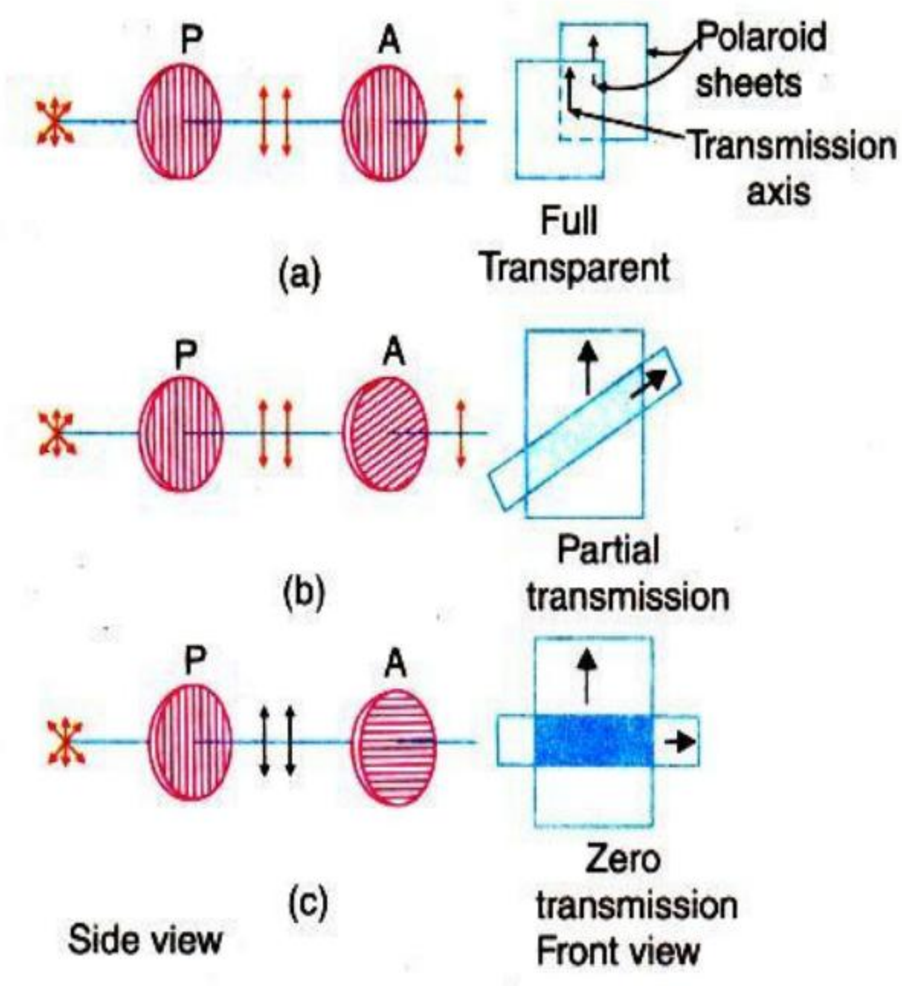


Figure : 14 (a), (b) & (c)

➢ Working

- When natural light is incident on the sheet, the component that is parallel to the chains of iodine atoms induces current in the conducting chains and is therefore strongly absorbed.
- Thus the transmitted light contains only the component that is perpendicular to the direction of molecular chains.
- The direction of E-vector in the transmitted beam corresponds to the transmission axis of the Polaroid sheet.

- These sheets are expensive and can be made in large sizes.
- They are widely used in sunglasses, camera filters etc. to eliminate the unwanted glare from objects.
- They can be used as polarizer and analyzer.
- Effect of Polarizer on natural light.

When unpolarized light passes through a polarizer, the intensity of the transmitted light will be exactly half that of the incident light. This can be proved as follows.

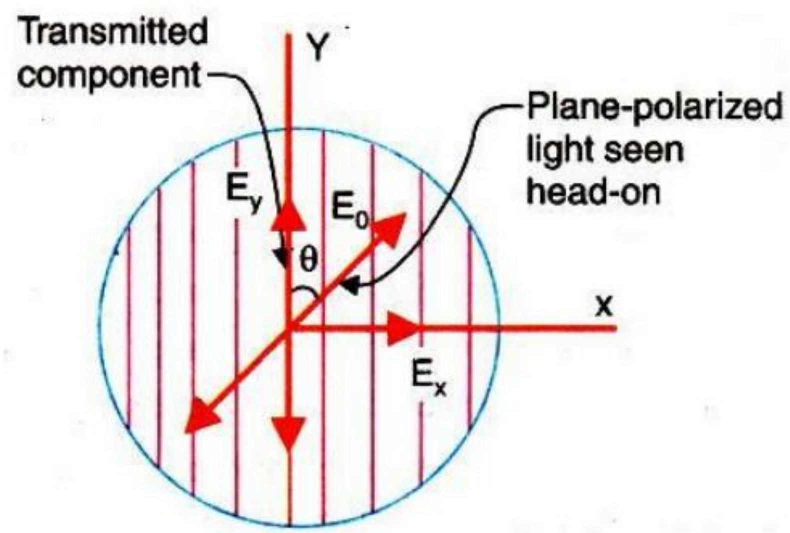


Figure : 15

Let E_0 is the amplitude of vibration of one of the waves of the unpolarized light incident on the polarizer.

Let E_0 makes an angle Q with the transmission axis of the polarizer.

Here E_0 may be resolved into its rectangular components E_x and E_y as shown in fig. The polarizer blocks the component E_x and transmits the component, Ey.

The intensity of the transmitted light is then,

$$I = E_y^2 (\cos\theta)^2$$

$$= E_0^2 (\cos^2\theta)$$

(Only Ey component of E0 is transmitted) But,

$$\text{Intensity} \propto (\text{Amplitude})^2$$

Therefore, $$I = I_0^2 (\cos^2\theta)$$

Now, the value Q is from 0 to 2x because unpolarized light has all possible vibrations.

$$I = I_0^2 \langle cos^2\theta \rangle$$

$$= \frac{I_0}{2\pi} \times \int_0^{2\pi} \cos 2\theta \, dx$$

$$= \frac{I_0}{2\pi} \times \int_0^{2\pi} \frac{1 + \cos 2\theta}{2} dx$$

$$I = \frac{I_0}{4\pi} \left[\int_0^{2\pi} dx + \int_0^{2\pi} \frac{1 + cos\ 2\theta}{2} dx \right]$$

$$I = \frac{I_0}{4\pi} \left([\theta]_0^{2\theta} + \left[\frac{sin\ 2\theta}{2} \right]_0^{2\pi} \right)$$

$$I = \frac{I_0}{4\pi} (2\pi + 0)$$

$$I = \frac{I_0}{2}$$

Thus, the intensity of transmitted light through the polarizer is half the intensity of incident light.

5.12 EFFECT OF ANALYZER ON PLANE POLARIZED LIGHT- MALUS LAW

When natural light is incident on a polarizer, the transmitted light is linearly polarized. If this light-further parses through an analyzer, the intensity varies with the angle between the transmission's axes of polarizer, and analyzer. Malus studied their phenomenon and known as Malus law.

5.12.1 Statement of Malus Law

It states that

"The intensity of the polarized light transmitted through the analyzer is proportional to cosine square of the angle between the plane of transmission of the analyzer and plane of transmission of the polarizer."

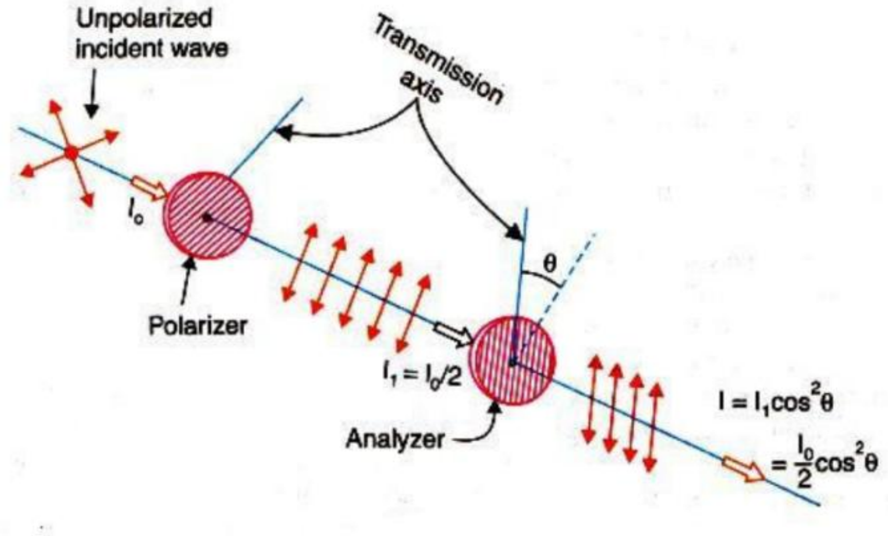


Figure : 16

Let I0 is the intensity of unpolarized light. The intensity of polarized light from the polarizer is I0/2. Take I1=I0/2.

This plane polarized light then passes through the analyzer. Let E is the aptitude of vibration and Q is the angle between this vibration and transmission axis of an analyzer. E resolves into two components (i) Ey, parallel to the plane of transmission of the plane of analyzer and (ii) Ex, perpendicular to the plane of analyzer.

Now, Ey component is only transmitted through the analyzer.

$$E_y = E\cos\theta$$

Intensity of light for this component:

$$I = E^2\, cos^2\,\theta = \, I_1\, cos^2\,\theta = \left(\frac{I_0}{2}\right)\, cos^2\,\theta$$

If

(i) $\theta = 0^0$, then axis are parallel I = I1

(ii) $\theta = 90^0$, then axis are perpendicular I= 0

(iii) $\theta = 180^0$, then axis are parallel I = I1

(iv) $\theta = 270^0$, then axis are perpendicular I = 0

Thus, there are two positions of maximum intensity and two positions of zero intensity when we rotate the axis of the analyzer with respect to that of the polarizer.

5.12 ANISOTROPIC CRYSTALS

Not all crystals are anisotropic in nature. Anisotropic is one of the properties exhibited by crystalline solids. The anisotropic property of a crystal depends on the symmetry of the unit cell in the crystal. The arrangement of these atoms in the crystal differs in all three planes. In anisotropic materials such as wood and composites, the properties vary along with the directions of the material.

- Diamond is crystalline and anisotropic, meaning that its properties are directional.
- Wood, composite materials, all crystals (except cubic crystal) are examples of anisotropic materials.
- Anisotropic crystals show birefringence, optical activity, dichroism and dispersion due to different refractive indices.

5.13.1 Isotropic Materials

In isotopic materials atoms are arranged in a regular periodic manner. In isotropic materials, when a light beam is incident, it refracts a single ray. It means that in such material the refractive index is same in all direction. e. g. Glass water and air.

5.13.2 Anisotropic Materials

In anisotropic material the arrangement of atoms differs in different directions within a crystal. Thus, the physical properties vary like, thermal conductivity, electrical conductivity, velocity of light and have refractive index etc. vary with the directions. Such crystal is then said Anisotropic.

The anisotropic crystals are divided into two classes.

(i) Uniaxial Crystal:

In this type of crystal, one of the refracted rays is on ordinary ray and the other is an extraordinary. e. g. Calcite, tourmaline and Quartz.

(ii) Biaxial Crystal:

In biaxial crystal both the refracted rays are extra ordinary rays.

e. g. mica, topaz & aragonite.

5.14 CALCITE CRYSTAL

Figure shows a calcite crystal in the form of a rhombohedron bounded by six parallelograms with angles equal to 78 and 102. At two opposite corners (A&H) the three angles of faces meeting there are all obtuse. These corners (A&H) the three angles of faces meeting there are all obtuse. These corners (A&H) are known as blunt corners.

- ### Optic axis

A line passing through 'A' making equal angles with each of the three corners gives the direction of optic axis. Any line parallel to this line is also an optic axis. In the figure AH is the optic axis of calcite crystal. If a ray of light is incident along the optic axis or in a direction parallel to the optic axis, then it will not split up into two rays. Thus the phenomenon of double refraction is absent when the light is allowed to enter the crystal along the optic axis.

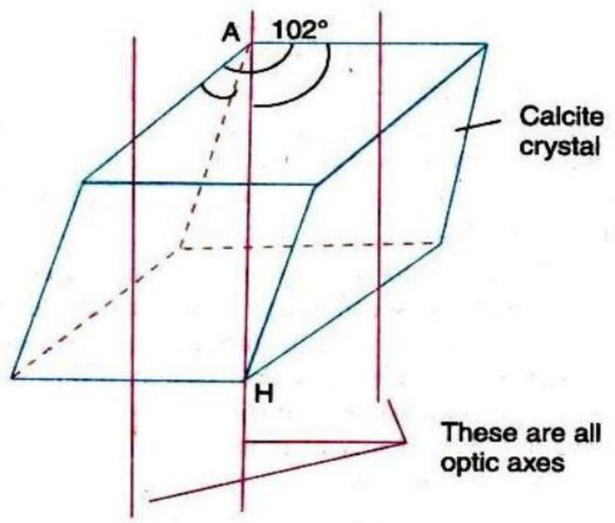


Figure : 17

➢ Principal Section

A plane containing the optic axis and perpendicular to a pair of opposite faces of the crystal is called the principal section of the crystal for that pair of faces.

As a crystal has six faces, so far every point inside the crystal there are three principal sections, one for each pair of opposite crystal faces. A principal section cuts the crystal surfaces in a parallelogram having angles 71º and 109º.

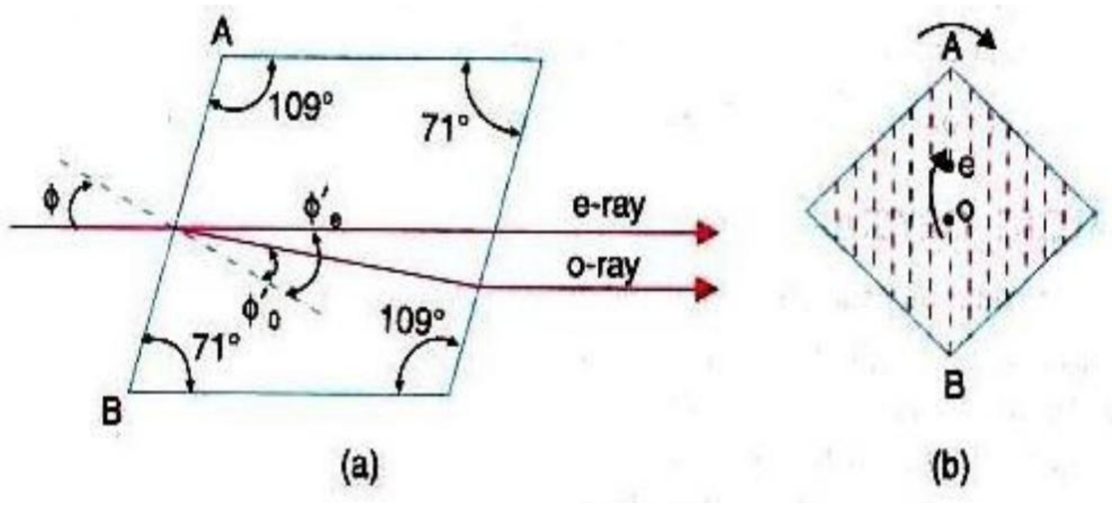


Figure : 18 (a) & (b)

In fig 18 (a) the principal section of the crystal is shown. An end view of any principal section is a straight line (shown by dotted line in fig. 18 (b))

The plane containing the optical axis and the O -ray is called the principal plane of O -ray. The plane containing the optic axis and the E-ray is called the principal plane of E-ray.

5.15 DOUBLE REFRACTION

When a ray of light is refracted by a crystal of calcite it gives two refracted rays. This phenomenon is called "DOUBLE REFRACTION".

The phenomenon of double refraction can be shown with the help of the following experiment.

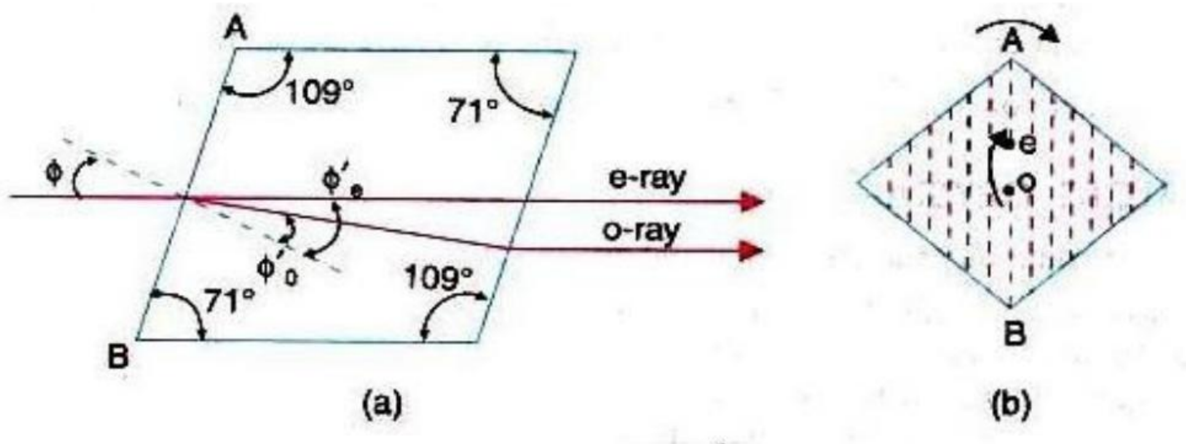


Figure : 19 (a) & (b)

Positive crystal: When reflective index for extraordinary ray is greater then that of O-ray $\mu e > \mu o$.

Negative crystal: when reflective index for extraordinary ray is lesser then that of O-ray $\mu e < \mu o$

Mark an ink dot on a piece of paper. If we place a calcite crystal over this dot, then two images of dots are observed. Now rotate the crystal slowly as shown in figure ii. It is found that our image remains stationary ink the second image rotates with the rotation of the crystal. The stationary image is known as the ordinary image while the second image is known as the extraordinary image. The retracted ray which produces ordinary image is known as ordinary ray O-ray and the retracted ray which produces extraordinary image is known extraordinary ray (E-ray).

When a ray of light AB is incident in the calcite crystal making our angle of incidence i, it is refracted along two paths inside the crystal

(i) Along BC making our angle of retraction r2 and

(ii) Along BD making our angle of refraction r1.

These two rays emerge out along CE and DO are parallel.

The difference between o-ray and e-ray is given below :

(1) The ordinary ray has a refractive index $\mu_0 = \frac{\sin i}{\sin r1}$ and the extraordinary ray has a refractive index $\mu_0 = \frac{\sin i}{\sin r2}$.

(2) The O-ray obeys the laws of refraction and its refractive index is constant. For E-ray its refractive index varies with the angle of incidence and its is not

fixed.

(3) For the case of calcite $\mu_0 > \mu_e$ because r1 less than r2. Therefore, the velocity of light for the O-ray inside the crystal is less than the velocity of light for E-ray. $\left(\mu_0 = \frac{c}{v^0} \text{ and } \mu_e = \frac{c}{v^e}\right)$.

(4) The O-ray travels in the crystal with same velocity in all directions ; where as the velocity of E-ray is different in different directions, because its refractive index varies.

(5) Both O-ray and E-ray is plane with the angle of incidence. Plane polarized they are polarized in mutually perpendicular planes.

5.15.1 Positive Crystal

In positive crystals the refractive index for E-ray is greater than refractive index for O-ray i.e. □e> □o.

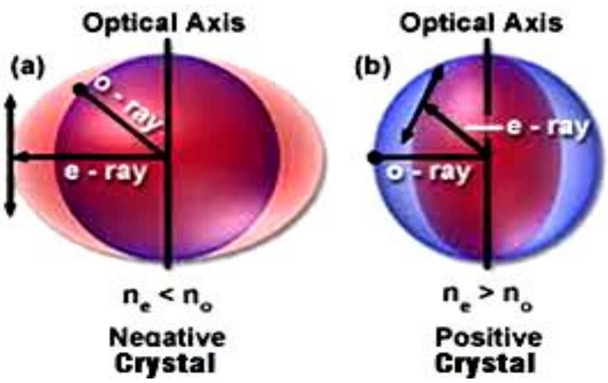
5.15.2 Negative Crystal

In negative crystals the refractive index for O-ray is greater than reflective index for e-ray i.e., □0> □e.

5.15.3 Difference Between Positive and Negative Crystal

Positive Crystal		Negative Crystal	
1.	In positive crystals the refractive index for E-ray is	1.	In negative crystals the refractive index for O-ray is

	greater than refractive index for O-ray i.e. □e> □o.		greater than reflective index for e-ray i.e., □0> □e.
2.	In positive crystals E-ray travels slower than o-ray in all directions except along the optic axis. V0> V0	2.	In negative crystals O-ray travels slower than E- ray in all directions except along the optic axis i.e. V0< V0
3.	According to Huygen's ellipse corresponding to e-ray is contained within the sphere corresponding to o-ray	3.	According to Huygen's ellipse corresponding to e-ray lies outside the sphere corresponding to O-ray
4.	Birefringence or amount of double refraction of a crystal is defined as Δ□=□e-.□ for Δ□ is positive quality or positive crystals	4.	Δ□ Negative for negative crystals.
5.	Example: Quartz	5.	Example: calcite



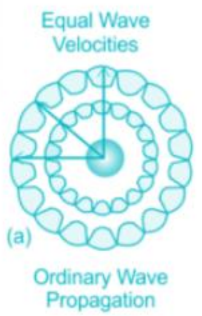
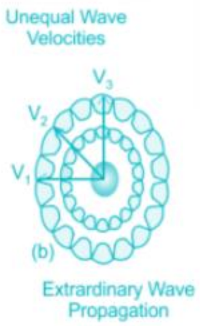
5.15.4 Ordinary Ray (O-Ray)

The part of a ray divided in two by double refraction that follows the ordinary laws of refraction because its speed is the same in all directions through the doubly refracting medium.

5.15.5 Extraordinary Ray (E-ray)

The part of a ray divided in two by double refraction that does not follow the ordinary laws of refraction because its speed varies with its direction in the doubly refracting medium.

5.15.6 Difference Between O-Ray and E-Ray

O-ray (Ordinary ray)	E-ray (Extraordinary ray)
It obeys Snell's law of refraction	It does not obey laws of refraction
It travels at the same speed in all direction inside the crystal	It travels at different speeds in different directions within the crystal. However, the speed of Ordinary and extraordinary ray is the same along its optic axis
The electric vector of O-ray vibrates perpendicular to the principal section of the O-ray	The electric vector of E-ray vibrates parallel to the principal section of the E-ray
	

5.16 HUYGEN'S EXPLANATION OF DOUBLE REFRACTION IN UNIAXIAL CRYSTAL

According to Huygen's they each point on a wave front act as a fresh source of disturbance and said secondary wavelets. He explained the phenomena of double refraction in Uniaxial crystal with the help of secondary wavelets.

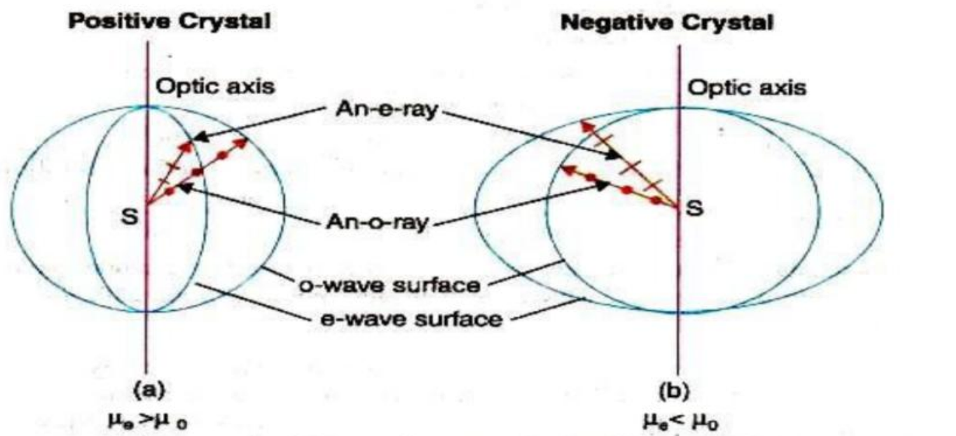
According to his theory:

1) When any wave front strikes a doubly refraction, crystal every point of the crystal becomes a course of two wave fronts.

2) Ordinary wave front corresponding to ordinary rays. Since ordinary rays have same velocity in all directions, the secondary wave front is spherical.

3) Extra-ordinary wave front corresponding to extra-ordinary rays. Since extra- ordinary rays have different velocities in different directions, the extra- ordinary wave front is ellipsoid of revolution, with optic axis as the axis of revolution.

4) The sphere and ellipsoid touch each other at points which lie on the optic axis of the crystal, because two velocity of ordinary and extra ordinary ray is same along the optic axis.



Huygens wave surfaces produced by a point source *S*-embedded in the birefringent crystal; (*a*) a positive crystal (*b*) a negative crystal.

Figure : 20 (a) & (b)

5) In certain crystals like calcite and tourmaline called the

negative crystal, the ellipsoid lies outside the sphere as shown in fig. 20 (a).

6) This shows that in negative crystals, the extra-ordinary wave front travels faster than ordinary wave front except along optic axis.

7) ($V_0 > V_0$ and $\mu_0 > \mu_e$).

8) In certain crystal (like quarter). Sphere lies outside the ellipsoid as shown in fig- 20 (b). Such crystals are called positive crystals. In the crystals, variety of ordinary wave front is greater than extraordinary wave front except along optic axis.

5.17 SUPERPOSITION OF WAVES LINEARLY POLARIZED AT RIGHT ANGLES

Let consider two light waves travelling in the x-direction one wave is polarized in x-y plane and the other is polarized y-z plane. Let us find effect produced because of the super positions of these two waves.

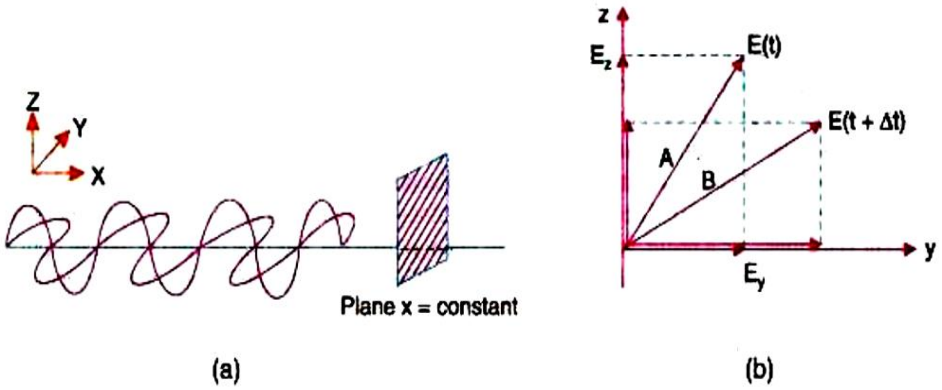


Figure : 21 (a) & (b)

Let, these two waves are represented by the following ways;

$E_y = E_1 \cos(kx - \omega t)$ (1)

$E_z = E_2 \cos(kx - \omega t + \delta)$ (2)

Where, δ = is phase difference between two waves

$u = \omega/2\pi$ = frequency

According to the principle of superposition,

$$E = E_y + E_z = E_1 \cos(kx - \omega t) + E_2 \cos(kx - \omega t + \delta) \quad (3)$$

From equation- (2),

$$E_2 = E_2 \, cos(kx - \omega t). cos\, \delta - E_2 \, sin(kx - \omega t + \delta)\, . sin\delta \quad (4)$$

$$= E_2 \, cos(kx - \omega t)\, . cos\, \delta \pm [1 - cos^2(kx - \omega t)]^{1/2} . E_2 \, sin\, \delta$$

Form equation – (1),

$$\cos(kx - \omega t) = \frac{E_y}{E_1}$$

$$E_2 = E_2 \frac{E_y}{E_1} cos\, \delta \pm \sqrt{1 - \frac{E_y^2}{E_1^2}} . E_2 \, sin\, \delta \quad (5)$$

Rearranging above equation,

$$\left[E_z - \frac{E_2}{E_1} E_y \cos \delta\right] = \pm \sqrt{1 - \left(\frac{E_y}{E_1}\right)^2} \,. E_2 \sin \delta$$

Squaring both the sides,

$$\left[E_z^2 + \left(\frac{E_2}{E_1} E_y\right)^2 \cos^2\delta\right] - \frac{2E_yE_zE_2}{E_1} \cos \delta$$
$$= E_2^2 \sin^2 \delta - \left(\frac{E_2}{E_1} E_y\right)^2 \sin^2\delta$$

$$E_2^2 + \left(\frac{E_2}{E_1} E_y\right)^2 (\cos^2\delta + \sin^2\delta) \cos^2\delta - \frac{2E_yE_zE_2}{E_1} \cos \delta$$
$$= E_2^2 \sin^2 \delta$$

$$E_2^2 + \left(\frac{E_2}{E_1} E_y\right)^2 - \frac{2E_yE_zE_2}{E_1} \cos \delta = E_2^2 \sin^2 \delta$$

(6)

Dividing both side by E_2^2 ,

$$\left(\frac{E_y}{E_1}\right)^2 + \left(\frac{E_z}{E_2}\right)^2 - \frac{2E_yE_zE_2}{E_1} \cos \delta = E_2^2 \sin^2 \delta$$

(7)

Above equation, is the general equation of ellipse, hence, the tip of the resultant vector traces an ellipse in Y-Z plane. The ellipse is constrained within a rectangle having sides 2E1, and 2E2.

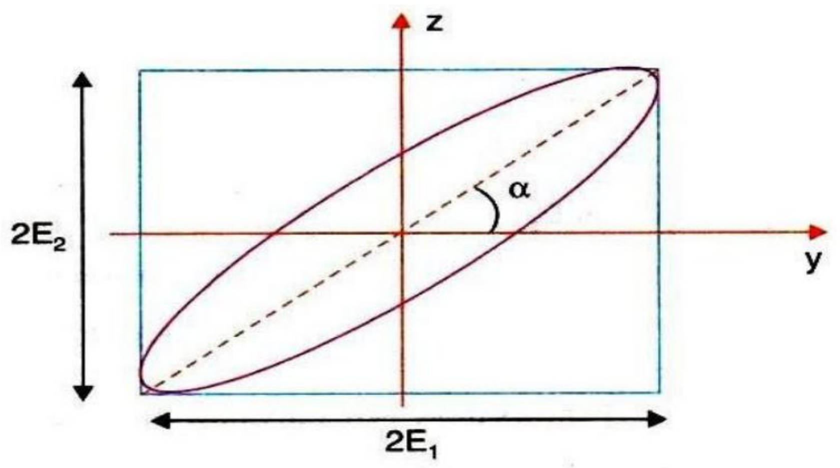


Figure : 22

➢ **Special cases :**

(1) When $\delta = 0$ or $\pm$ 2mx, then two waves are in phase.

Therefore,

$\cos\delta = 1\ and \sin\delta = 0$

Then, equation (4) becomes,

$$\left(\frac{E_y}{E_1}\right)^2 + \left(\frac{E_z}{E_2}\right)^2 - \frac{2E_yE_zE_2}{E_1} = 0$$

$$\left[\frac{E_y}{E_1} - \frac{E_z}{E_2}\right]^2 = 0$$

$$\frac{E_y}{E_1} - \frac{E_z}{E_2} = 0$$

$$E_2 = \frac{E_2E_y}{E_1}$$

(8)

The above equation represents a straight line, having a slope (E2/E1). It means that, the resultant of two plane-polarized waves is again a plane-polarized wave.

(2) When $\delta = \pi \; or \; \pm \; (2mx + 1)\pi$

The two waves are in opposite phase.

Therefore,

$$cos\, \delta = -1 \quad and\; sin\, \delta = 0$$

Then, equation (4) reduced to,

$$\left(\frac{E_y}{E_1}\right)^2 + \left(\frac{E_z}{E_2}\right)^2 - \frac{2E_yE_zE_2}{E_1} \; = \; 0$$

$$\left[\frac{E_y}{E_1} - \frac{E_z}{E_2}\right]^2 = 0$$

$$\frac{E_y}{E_1} - \frac{E_z}{E_2} = 0$$

$$E_2 = \; \frac{E_2E_y}{E_1}$$

(9)

This equation represents a straight line of a slope (-E2/E1). It means that the resultant of two plane polarized wares, which are in opposite phase, is against a plane-polarized wave.

(3) If $\delta = \frac{\pi}{2} \; or \; \pm (2mx + 1)\, \pi/2$

$$cos\, \delta = 0 \quad and\; sin\, \delta = 1$$

Then, equation (4) reduced to,

$$\frac{E_y}{E_1} - \frac{E_z}{E_2} = 1$$

(10)

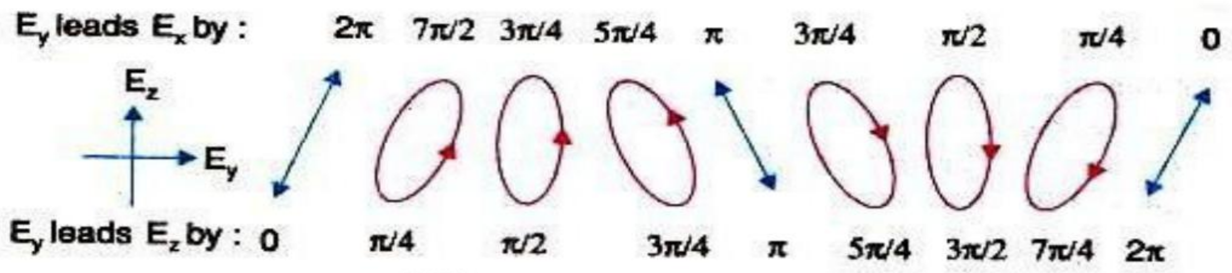
This is the equation of ellipse. Its major and minor axis considers with y-and z coordinates axes. Thus the waves are out of phase by 90^0 and their resultant wave is elliptically polarized wave.

(4) If $\delta = \frac{\pi}{2}$ and $E_1 = E_2 = E_0$

Then, equation (4) reduced to,

$E_1^2 + E_2^2 = E_0^2$

This is the equation of circle. Here result wave is circularly polarized.



Various polarization configurations.

5.18 TYPES OF POLARIZED LIGHT

There are different types of Polarized Light :

5.18.1 Unpolarized Light

It consists of sequence of wave trains, all oriented at random. It is considered as the resultant of two optical vectors. Components, which are incoherent.

5.18.2 Linearly Polarized Light

It can be regarded as a resultant of two coherent linearly polarized waves.

5.18.3 Partially Polarized Light

It is a mixture of linearly polarized light and unpolarized light.

5.18.4 Elliptically Polarized Light

It is the resultant of two coherent waves having different amplitudes and a constant phase difference of 90^0. In elliptically polarized light, the magnitude of electric vector E rotates about the direction of propagation.

If light is coming towards us, we would observe that tip of the E vector traces an ellipse.

- The side view of E vector gives flattered helix in space.
- If we look from the source and rotation of E vector is clockwise then it is right elliptically polarized wave. If it is anti clock wise then it is left elliptically polarized wave.

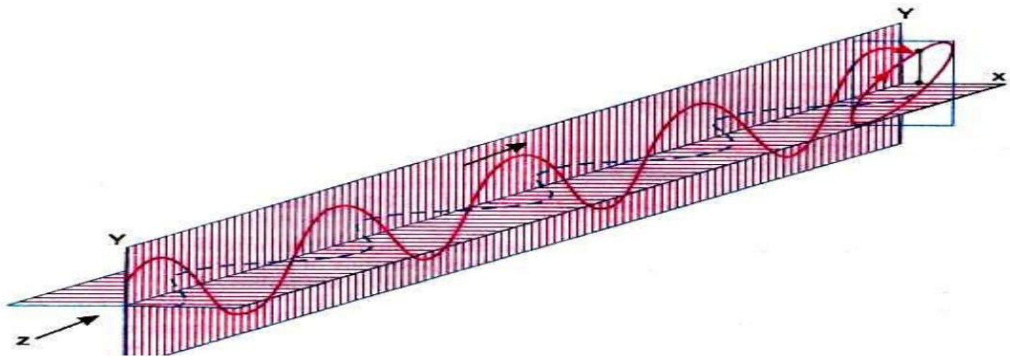


Figure : 23

5.18.5 Circularly Polarized Lsight

- It is the resultant of two coherent waves having same amplitudes and a constant phase difference of 90^{o}. In this type of light the magnitude of E vector remains constant.
- If light is coming towards us, we would observe that tip of the E vector trances a circle.
- If we look from the source and rotation E vector is clockwise then it is said right. Circularly polarized and if it anticlockwise then it is left circularly polarized wave.

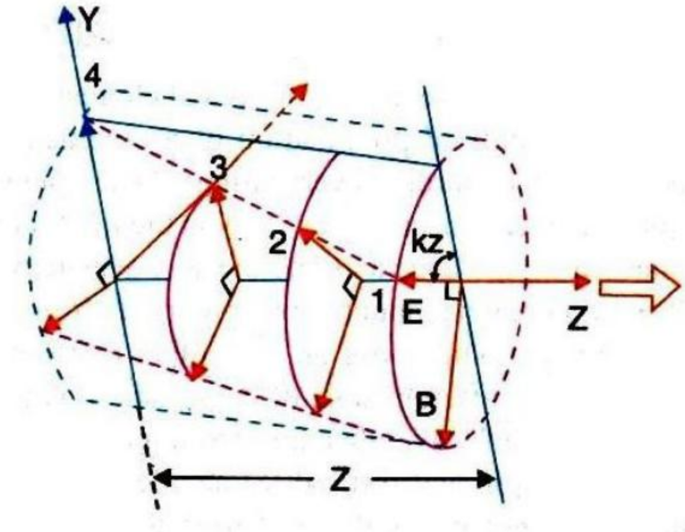


Figure : 24

5.19 LIQUID CRYSTAL DISPLAY (LCD)

Liquid crystal Display is most widely used device which makes the use of polarization. It is used in wrist watches, educators, clocks, electronic instruments, video games etc.

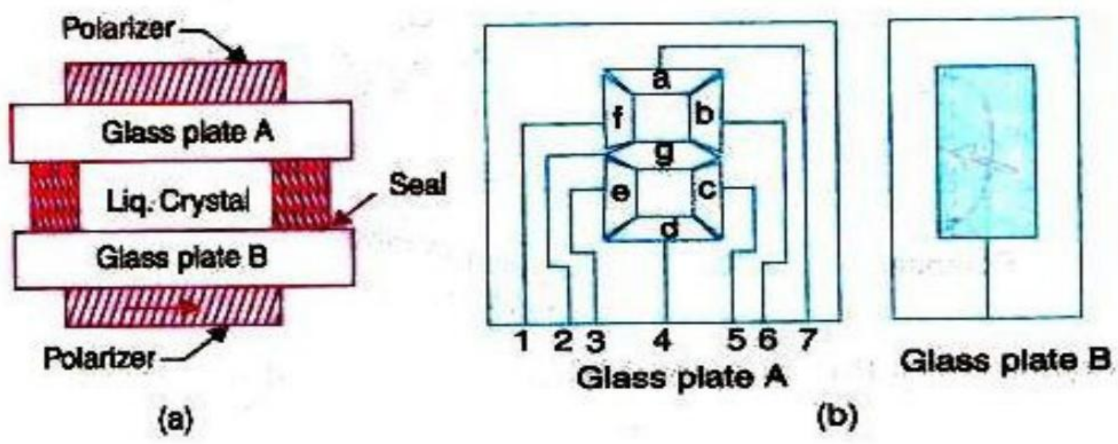
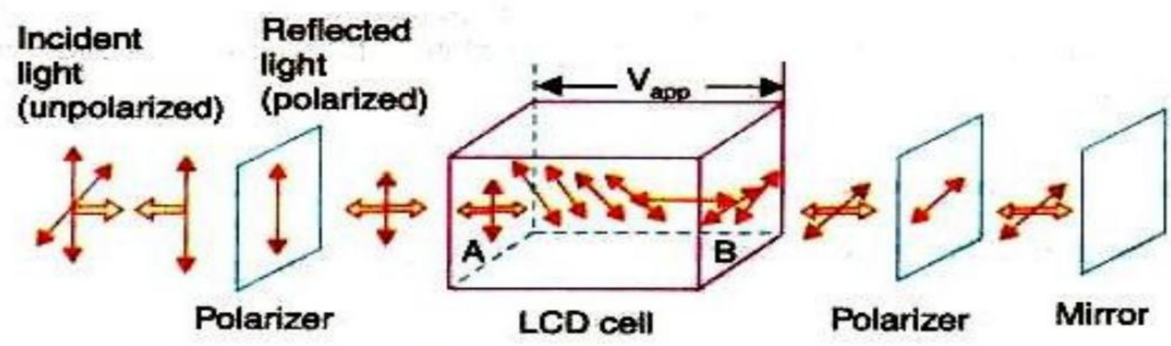


Figure : 25

5.19.1 Construction of LCD

- An LCD consists of liquid crystal material of 10□m thickness.
- It is double refracting material.
- This material is supported between thing glass plates.
- The inner surfaces of thin glass are coated with transparent conducting material.
- This conducting material is etched in the form of a digit or character as shown in fig.2.
- The assembly of glass plates with liquid crystal material is sandwiched between two crossed polarizer sheets.

5.19.2 Working of LCD



The action of an LCD having a twisted molecular arrangement.

Figure : 26

- During fabrication of LCD, the liquid crystal molecules are arranged as shown in fig.3.
- This arrangement of molecules is called twisted molecular arrangement i.e. 90^0. Rotation from plate A to B.
- When natural light is incident on the LCD, the front polarizer converts it into linearly polarized light.
- When this polarized light propagates through LCD, the optical vector is rotated by 90^0 because of twisted molecular arrangement.
- This light passes very easily through the rear polarizer whose transmission axis is perpendicular to that of the front polarizer.
- A reflecting coating at the back of the rear polarizer sends back the light, which comes out from the front polarizer.
- The display seems illuminated uniformly.

- When a voltage is applied to the device, the molecules between the electrodes align along the directions of field.
- When light passes through this region optical vector does not undergo rotation.
- The rear polarizer blocks the light and therefore a dark digit or character is seen in that region as shown in figure.

5.20 SUMMARY

Polarization of light is a property that applies to turning waves that shows the geometrical blooming of the oscillations. In a turning wave, the way of the oscillation is ninety degrees of the motion of the wave. Plane polarized light has the two waves in which the way of vibration is similar for all waves. In circular polarization, the electric vector turns about the way of straight light as the wave progresses. If you glow a beam of polarised moonlight that is chromated light (light of only the one frequency – in different words a similar colour) through a solution of a metamorphic active substance, the light that comes out, its plane of polarisation is seen to have rotating or turning around. The rotating body may be clockwise or anti-clockwise.

5.21 TERMINAL QUESTIONS

1. What do you mean by polarization? Is it possible in longitudinal waves?

2. Distinguish between polarized and unpolarized light.

3. Explain the terms plane of polarization and plane of vibration. What is the angle between them?

4. State Bresters’s law. How is refractive index determined using this law?

5. State and explain Malus law.

6. Give assumptions of Hugens’ theory to explain the phenomenon of double refraction.

7. Describe the construction and working of a Nicol prism.

8. State Brewster’s law and use it to prove that when light is incident on a transparent substance at the polarizing angle, the reflected and refracted rays are at right angles to each other.

9. Explain the action of pile of plates in producing plane polarized light.

10. Distinguish between positive and negative crystals. Describe the phenomenon of double refraction in uniaxial crystals. How is double refraction explained by Huygens’ theory?

11. Explain on the basis of Huygens’ theory, the propagation of light through a uniaxial crystal.

12. Polarization of light conclusively proves that

    (a) light waves are longitudinal

    (b) light waves are transverse

    (c) light waves and sometimes longitudinal and sometimes transverse

    (d) light waves are mechanical like sound waves

13. Which one of the following phenomenon decides that light waves are transverse

(a) Reflection

(b) Polarization

(c) Double refraction

(d) Interference

14. Propagation of polarized light is

(a) one dimensional

(b) two dimensional

(c) three dimensional

(d) similar to unpolarized light

15. A calcite crystals is

(a) uniaxial crystal

(b) biaxial crystal

(c) opaque crystal

(d) multiaxial crystal

16. A calcite crystal is

(a) positive crystal

(b) negative crystal

(c) it cannot produce polarized light

(d) it has the following wavefront: Ellipsoid inside sphere

17. A Nicol prism is made of

(a) glass

(b) calcite

(c) quartz

(d) plastic

18. A Nicol prism can be used

(a) as a polarizer only

(b) as an analser only

(c) both (a) and (b)

(d) neither (a) nor (b)

19. There is no role of magnitude of wavelength of light in the phenomenon of

(a) interference

(b) diffraction

(c) polarization

(d) all above

20. The plane of vibration and plane of polarization are

(a) parallel

(b) perpendicular

(c) inclined of 45^o with each other

(d) inclined arbitrarily

5.22 SOLUTION AND ANSWER OF TERMINAL QUESTION

1. Section 5.3
2. Section 5.18
3. Section 5.4.2 and 5.4.3
4. Section 5.6.2 and 5.6.3
5. Section 5.12.1
6. Section 5.16
7. Section 5.11.2
8. Section 5.6.2
9. Section 5.7.1
10. Section 5.15.3
11. Section 5.16

12. (b) 13. (b) 14. (b) 15. (a) 16. (b) 17. (b)

18. (c) 19. (c) 20. (b)

5.23 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.

2. Optics – Ajoy Ghatak

3. Optics – Principles and Application, K. K. Sharma, Academic Press.

4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.

5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT 06: DETECTION OF POLARIZED LIGHT

STRUCTURE:

6.1 INTRODUCTION

In the previous unit we discussed about the concept of Polarization of Light. In the present unit, we explore Polarization in more detail, presenting different representation of polarization properties as well as extending the notion of polarization properties beyond that of linear polarization.

Sunlight and almost every other form of natural artificial illumination produced light waves whose electric field vectors vibrate in all planes that are perpendicular with respect to the direction of propagation. If the electric field vectors are restricted to a single plane by filtration of the beam with specialized materials, then the light is referred to as plane or linearly polarized with respect to the direction of propagation, and all waves vibrating in a single plane are termed plane parallel or plane-polarized.

Also, in this unit we shall study about Babinet compensator, optical Rotation and specific Rotation.

Lastly, we also understand the concept of Polarimeters.

6.2 OBJECTIVES

After studying this unit, you should be able to :

- Understand the types of Polarization of Light.
- Distinguish among plane, Circular and Elliptical Polarized Light.
- Define Retardation Plates.
- Describe the Concept of Quarter and Half Wave Plate.
- Understand the concept of Babinet Compensator.
- Explain the Idea of Polarimeters.

6.3 WHAT IS LIGHT DETECTION ?

Light detection and ranging (LIDAR) refer to **an optical system that measures any or all of a variety of target parameters including range, velocity, and chemical constituents.**

Light Detectors are **used to simply detect light**. Within this circuit, whenever you place your finger on top of the light dependent resistor in a room that contains light, the LED will turn off.

A device used in an optical transmission system to detect an optical signal generated by a light source and propagating through a medium. A light detector essentially is an optical receiver that is paired with an optical transmitter, both of which are connected to electrically based devices or systems.

6.3.1 Plane Polarization Light

Polarization is a characteristic of transverse waves that determines the geometrical orientation of their oscillations. A basic example of a polarized transverse wave is vibrations moving along a guitar string (taut string). Depending on how strings are plucked, the vibration can be in a horizontal direction, vertical direction, or at any angle perpendicular to the strings. On the other hand, in longitudinal waves, like sound waves in a gas or liquid, the particles' displacement in the oscillation is every time in the propagation's direction. Therefore, these waves do not show polarization. Transverse waves that show polarization include light, radio waves, shear waves, etc.

If by filtering the beam with specialized materials, the electric field vectors are limited to a single plane, then the light is referred to as plane or linearly polarized with respect to the propagation direction. All waves vibrating in a single plane are referred to as plane parallel or plane-polarized.

Polarization is a science-based phenomenon that encompasses light, radiation, or magnetism travelling in particular directions. Polarization generally refers to how people think outside of science, especially as two viewpoints arise that pull people apart, sort of like two competing magnets.

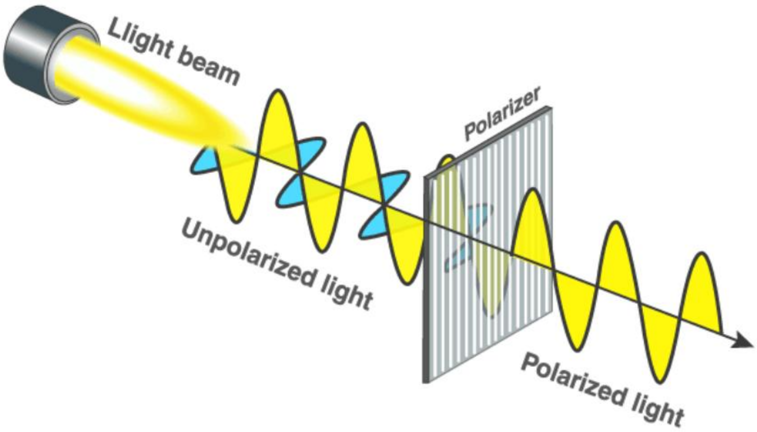


Figure : 1

6.3.2 Circular Polarized Light

Circularly polarized light is made up of dual perpendicular electromagnetic plane waves of identical amplitude and 90° variation in phase.

In the field of electrodynamics, the direction and the strength of an electric field are represented by the electric field vector.

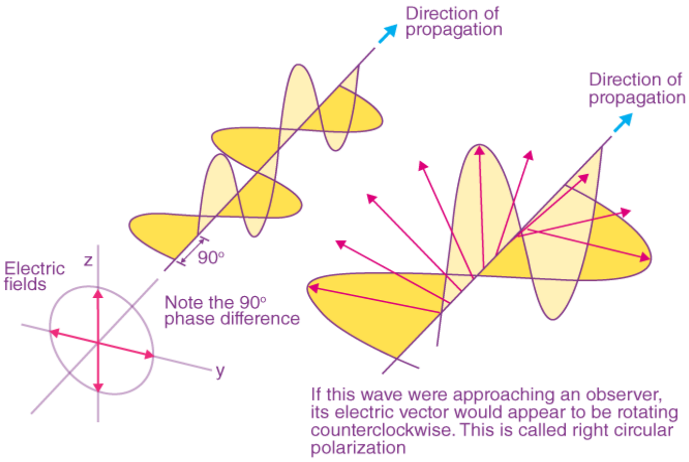


Figure : 2

In the instance of a circularly polarized wave, the tip of the field vector, at a particular point in space, relates to the light's phase as it propagates through space and time. At any point in time, the electric vector of the wave shows a point on a helix aligned along the propagation's direction. A circularly polarized wave can turn into one of two attainable senses. One is right-handed or clockwise circular polarization in which the field vector turns in a right-hand sense relative to the propagation's direction, and left-handed or counter-clockwise circular polarization in which the electric field vector turns in a left-hand sense. Circular polarization is a special case of elliptical polarization

6.3.3 Elliptical Polarized Light

The electric field of light follows an elliptical propagation. The amplitude and phase difference between the two linear components are not equal.

In electrodynamics, elliptical polarization is the polarization of electromagnetic radiation such that the tip of the electric field vector describes an ellipse in any fixed plane intersecting, and normal to, the direction of propagation. An elliptically polarized wave may be resolved into two linearly polarized waves in phase quadrature, with their polarization planes at right angles to each other. Since the electric field can rotate clockwise or counter clockwise as it propagates, elliptically polarized waves exhibit chirality.

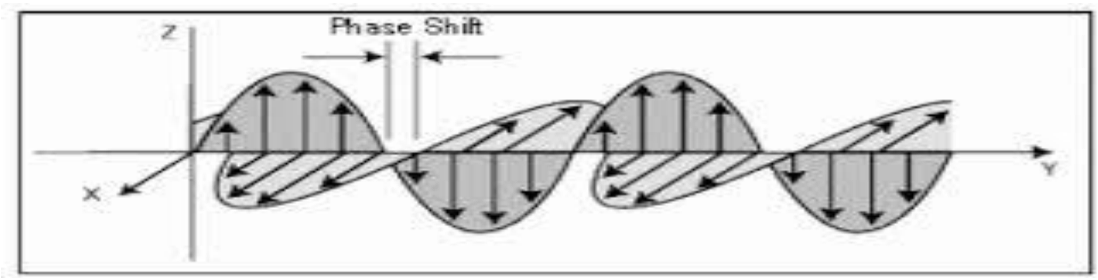


Figure : 3

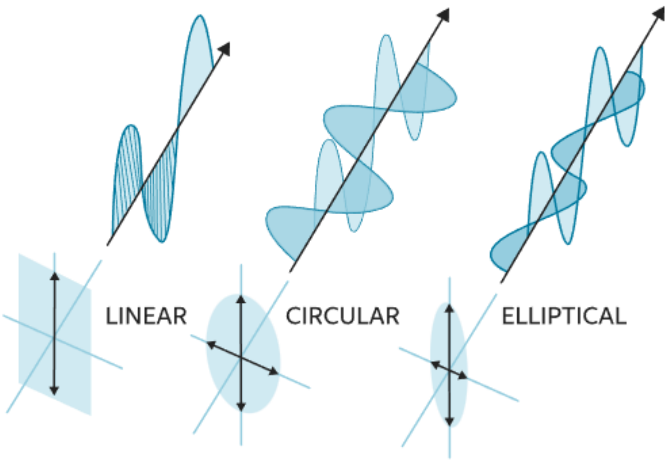
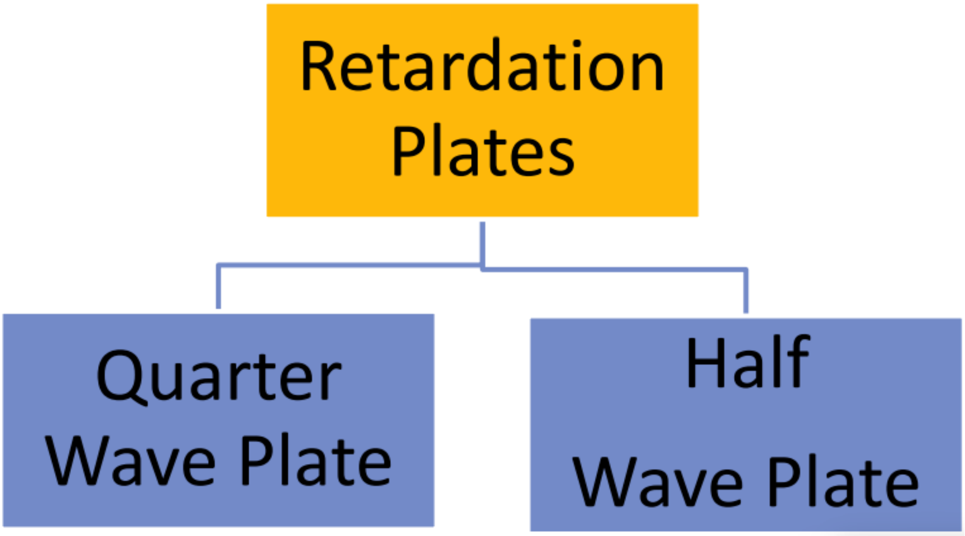


Figure : 4

6.4 WHAT IS RETARDATION PLATE

Retardation plates, also referred to as wave plates, produce a phase shift, the principle of which is based on polarization.

As we have already discussed that when O-ray and E-ray are emerging from the crystals, they have a path difference between them due to variation in their velocities. These crystals are known as retardation plates. These are mainly of two types.



6.4.1 Quarter Wave Plates

A plate of doubly refracting uniaxial crystal (quartz or calcite) whose refracting faces are cut parallel to the direction of optic axis and whose thickness is such as to produce a phase difference of $\frac{\pi}{2}$ or a path difference of $\frac{\lambda}{4}$ between the ordinary and extraordinary rays is called quarter wave plate.

Let d be the thickness of the quarter wave plate and n_o and n_e be the refractive indices of the medium for the ordinary and extraordinary rays respectively. For normal incidence, the path difference between the extraordinary and the ordinary rays is given by

$path\ difference = d\,(n_o - n_e)$ for the negative crystals and

$path\ difference = d\,(n_e - n_o)$ for the positive crystals.

But the path difference for a quarter wave plate is equal to $\frac{\lambda}{4}$.

Thus for negative crystals $d\,(n_o - n_e) = \frac{\lambda}{4}$ or $\boldsymbol{d = \frac{\lambda}{4(n_o - n_e)}}$

Similarly for the positive crystals $d\,(n_e - n_o) = \frac{\lambda}{4}$ or $\boldsymbol{d = \frac{\lambda}{4(n_e - n_o)}}$

A quarter wave plate introduces a phase difference of $\Delta\phi$ between E -O- ray and O- ray given by $\Delta\phi = \frac{2\pi}{\lambda} \times \Delta x = \frac{2\pi}{\lambda} \times \frac{\lambda}{4} = \frac{\pi}{2} = 90^0$.

If the thickness of the plate is such that $d\,(n_o - n_e) = (2m+1)\frac{\lambda}{4}$ where $m = 0, 1, 2, ...$ Then the plate still acts as a quarter wave plate for all these integers.

A quarter wave plate is used to produce and detect circularly and elliptically polarized light. If the plane polarized light incident on the quarter wave plate with its vibrations making an angle of 45^o with the optic axis, then the emergent light is circularly polarized. For angles other than 45^o, it is elliptically polarized.

6.4.2 Half Wave Plates

A plate of doubly refracting uniaxial crystal (quartz or calcite) whose refracting faces are cut parallel to the direction of optic axis and whose thickness is such as to produce a phase difference of π or a path difference of $\frac{\lambda}{2}$ between the ordinary and extraordinary rays is called quarter wave plate.

Let d be the thickness of the half wave plate and n_o and n_e be the refractive indices of the medium for the ordinary and extraordinary rays respectively. For normal incidence, the path difference between the extraordinary and the ordinary rays is given by

$path\ difference = d\,(n_o - n_e)$ for the negative crystals and

$path\ difference = d\,(n_e - n_o)$ for the positive crystals.

But the path difference for a half wave plate is equal to $\frac{\lambda}{2}$.

Thus for negative crystals $d\,(n_o - n_e) = \frac{\lambda}{2}$ or $\boldsymbol{d = \frac{\lambda}{2(n_o - n_e)}}$

Similarly for the positive crystals $d\,(n_e - n_o) = \frac{\lambda}{2}$ or $\boldsymbol{d = \frac{\lambda}{2(n_e - n_o)}}$

A half wave plate introduces a phase difference of $\Delta\phi$ between E -O- ray and O- ray given by $\Delta\phi = \frac{2\pi}{\lambda} \times \Delta x = \frac{2\pi}{\lambda} \times \frac{\lambda}{2} = \pi = 180^0$.

If the thickness of the plate is such that $d\,(n_o - n_e) = (2m+1)\frac{\lambda}{2}$ where $m = 0, 1, 2, ...$ Then the plate still acts as a half wave plate for all these integers.

When a incident plane polarized light is incident on a half wave plate, the emergent light is also plane polarized for all orientations of the plate with respect to the plane of vibration of incident light.

Quarter wave plate and the half wave plate are called retarding plates as they retard one of the beams.

6.5. PRODUCTION AND DETECTION OF PLANE POLARIZED LIGHT

Plane polarized light

Production : A unpolarised light is allowed pass through a Nicol prism. The light is split in to ordinary and extraordinary rays. The ordinary ray is totally reflected by the Canada balssm material in the Nicol prism. Only extraordinary ray with vibrations parallel to principal section of Nicol prism emerges which is plane polarized light.

Detection : Plane Polarised Light beam is allowed to fall on Nicol prism. The Nicol prism is rotated. The intensity of emitted light gradually decrease and become zero at two positions in each rotation. That is, the intensity varies from zero to maximum. Then the incident light is said to be plane polarized.

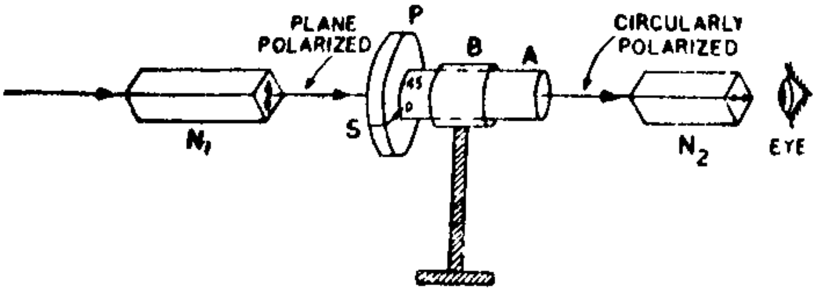
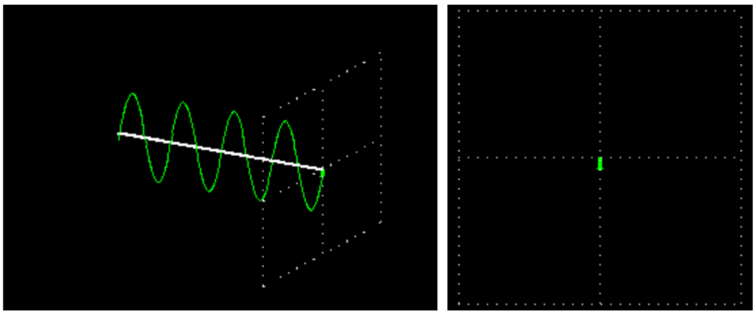


Figure : 5

Plane-polarized light

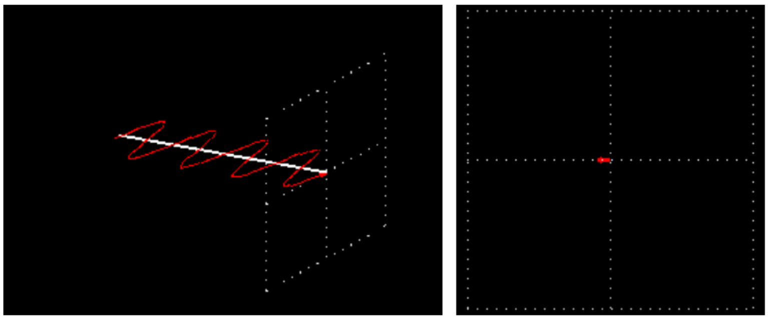
Vertical

$$E_y = A\sin(x/\lambda - \omega t)$$



Horizontal

$$E_z = A\sin(x/\lambda - \omega t)$$



6.6 PRODUCTIN AND DETECTION OF ELLIPTICAL POLARIZED LIGHT

Elliptically polarized light: **Production :**

1. To produce elliptically polarized light, the two waves vibrating at right angles to each other having the unequal amplitude should have a phase difference of $\frac{\pi}{2}$ or a path difference of $\lambda/4$.
2. The experimental arrangement is same as above. A beam of monochromatic light falls on the Nicol prism N_1 . The emergent light is plane polarized.
3. When another Nicol prism N_2 is placed at a suitable distance in the path of polarized light, It is rotated till the field of view is dark. Now the two nicols are said to be crossed.
4. Now a quarter wave plate is introduced between the prisms. The field of view is not dark. The quarter wave plate is rotated so that the field of view becomes dark. At this position it is observed that polarized light falling on the plate has its vibrations parallel to optic axis of the plate and perpendicular to N_2.
5. Now the quarter wave plate is rotated through an angle other than 45^0 so that the vibrations of light falling on the plate make an angle other than 45^0 with the optic axis. Now the amplitudes of vibrations of the two rays are unequal and there is a phase difference of $\frac{\pi}{2}$ between them. This results in elliptically polarized light.

Detection :

1. The light beam is allowed to fall on Nicol prism. If on rotation of Nicol prism, the intensity of emitted light varies from maximum to minimum, then light is either elliptically polarised or a mixture of plane polarized and unpolarised.
2. To differentiate between the two, the light is first passed through quarter wave plate and then through Nicol prism.
3. If beam is elliptically polarised, then after passing through quarter wave plate, an extra path difference of $\lambda/4$ is introduced between O-ray and E-ray and get converted into plane polarized
4. Thus, on rotating the Nicol, the light can be extinguished. If, on the other hand, beam is mixture of polarised and unpolarised it remains mixture after passing through quarter wave plate and on rotating the Nicol intensity of emitted light varies from maximum to minimum.

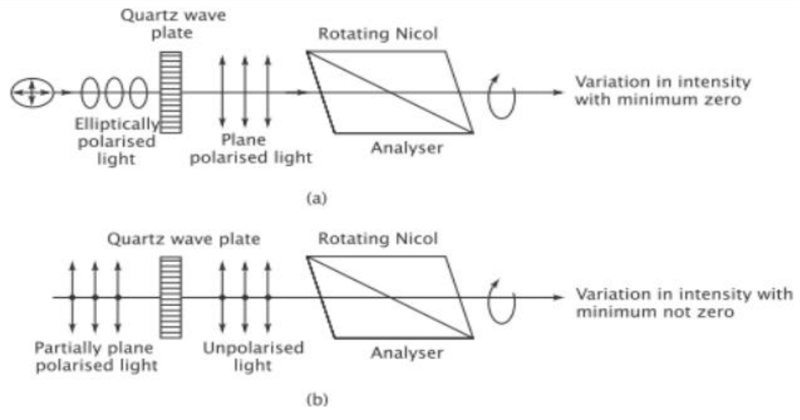


Figure : 6

Elliptical polarization

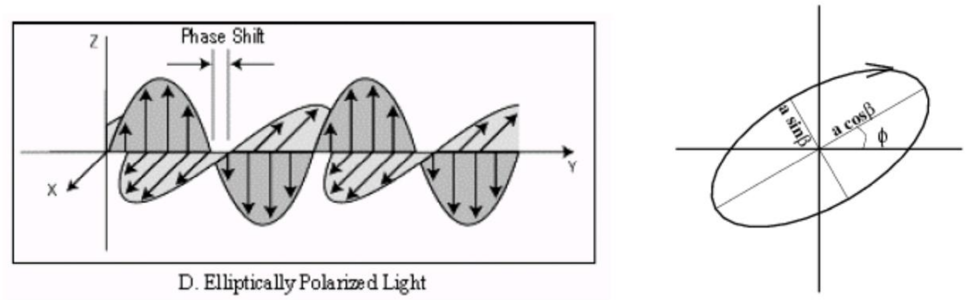


Figure : 7

Linear + circular polarization = elliptical polarization

6.7 PRODUCTION AND DETECTON CIRCULARLY POLARIZED LIGHT

Production: To produce circularly polarized light, the two waves vibrating at right angles to each other having the same amplitude and time period should have a phase difference of $\pi 2$ or a path difference of $\lambda/4$.

1. The experimental arrangement is as shown. A beam of monochromatic light falls on the Nicol prism N_1 . The emergent light is plane polarized.
2. When another Nicol prism N_2 is placed at a suitable distance in the path of polarized light, It is rotated till the field of view is dark. Now the two nicols are crossed.
3. Now a quarter wave plate is introduced between the prisms. The field of view is not dark. The quarter wave plate is rotated so that the field of view becomes dark. At this position it is observed that polarized light falling on the plate has its vibrations parallel to optic axis of the plate and perpendicular to N_2.
4. Now the quarter wave plate is rotated through 45^0 so that the vibrations of light falling on the plate make an angle of 45^0 with the optic axis. Now the amplitudes of vibrations of the two rays are equal and there is a phase difference of $\frac{\pi}{2}$ between them. This results in circularly polarized light.

Detection :

1. The light beam is allowed to fall on a Nicol prism. If on rotation of Nicol prism the intensity of emitted light remains same, then light is either circularly polarised or unpolarised.
2. To differentiate between unpolarised and circularly polarised light, the light is first passed through quarter wave plate and then through Nicol prism.

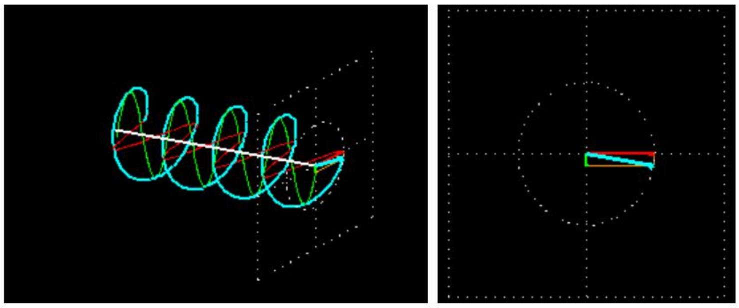
3. If beam is circularly polarised then after passing through quarter wave-plate an extra difference of λ/4 is introduced between ordinary and extraordinary component and gets converted into plane polarised. Thus on rotating the Nicol, the light can.be extinguished at two plates.

4. If, on the other hand, the beam is unpolarised, it remains unpolarised after passing through quarter wave plate and on rotating the Nicol, there is no change in intensity of emitted light.

Right circular

$$E_y = A\sin(x/\lambda - \omega t + 90°)$$

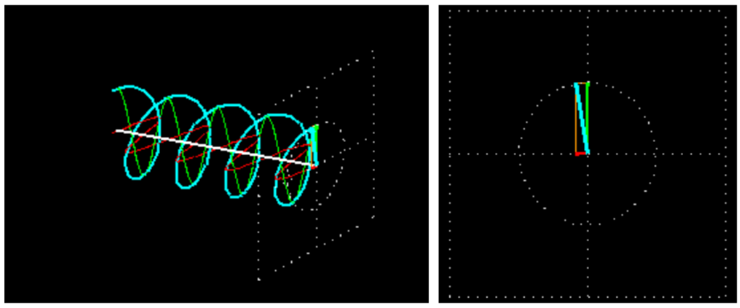
$$E_z = A\sin(x/\lambda - \omega t)$$



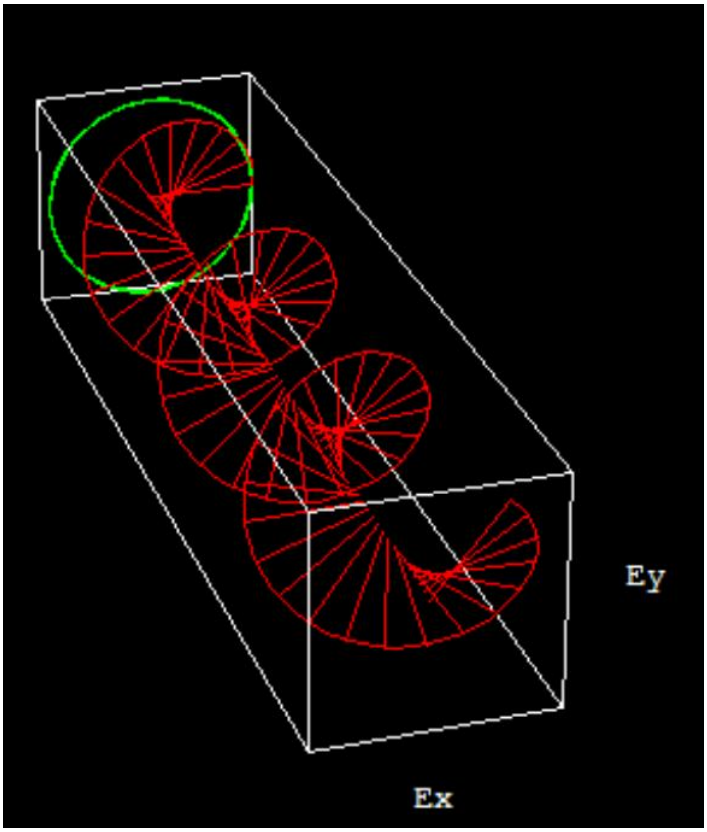
Left circular

$$E_y = A\sin(x/\lambda - \omega t - 90°)$$

$$E_z = A\sin(x/\lambda - \omega t)$$



Circular polarization



Ey
Ex

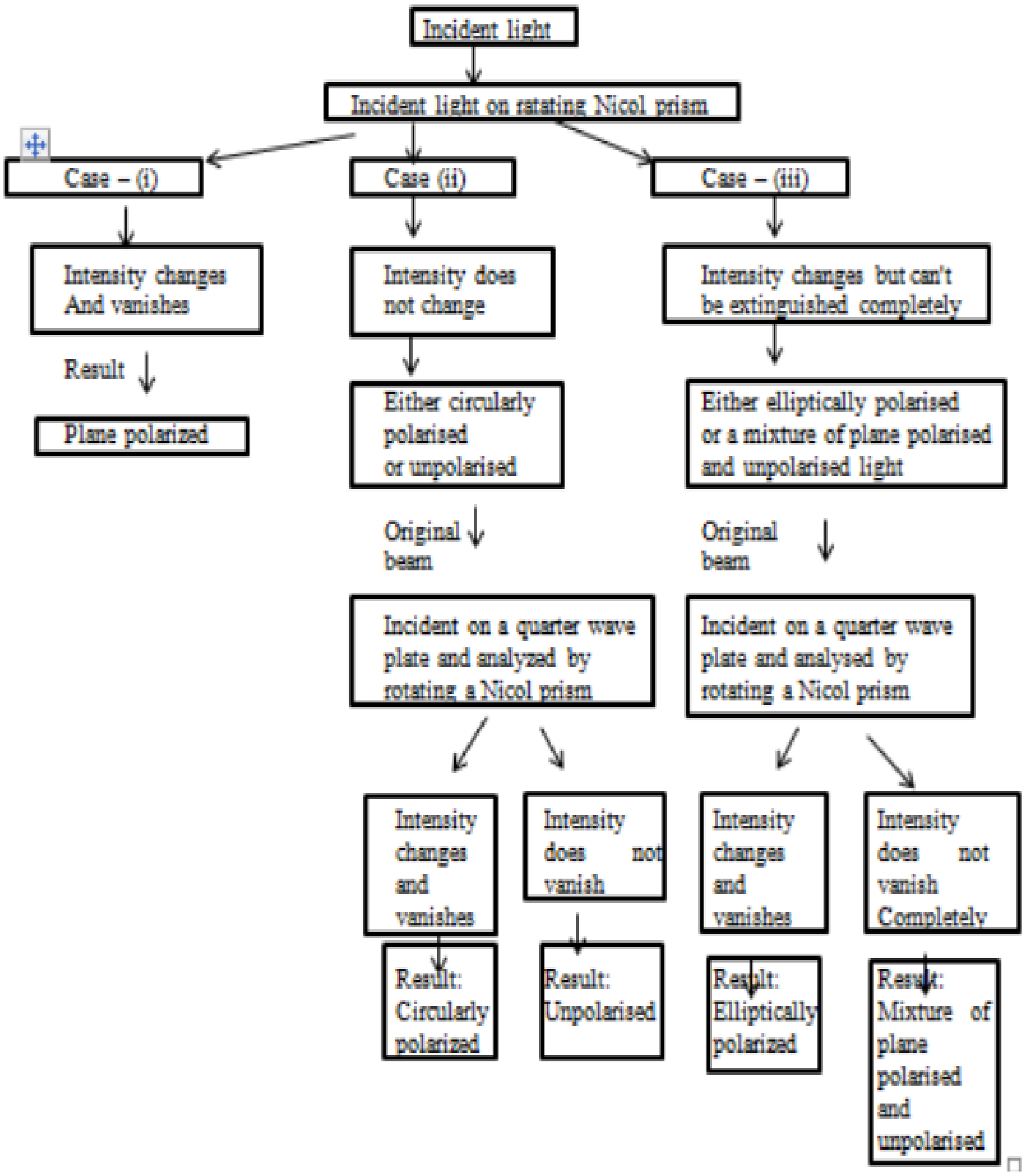


Figure : 8

6.8 BABINET COMPENSATOR

The Babinet–Soleil compensator is **a continuously variable, zero-order retarder**. It consists of two birefringent wedges, one of which is movable, and another is fixed to a compensator plate. The orientation of the long axis of the wedges is perpendicular to the long axis of the compensator plate.

6.8.1 Principle of Babinet Compensator

We know that for a given wavelength, a quarter wave plate and half wave plate produces a fixed path difference of $\frac{\lambda}{4}$ and $\frac{\lambda}{2}$ respectively between ordinary and extraordinary ray. It is useful for light of a particular wavelength and different plates are to be used for different wavelengths.

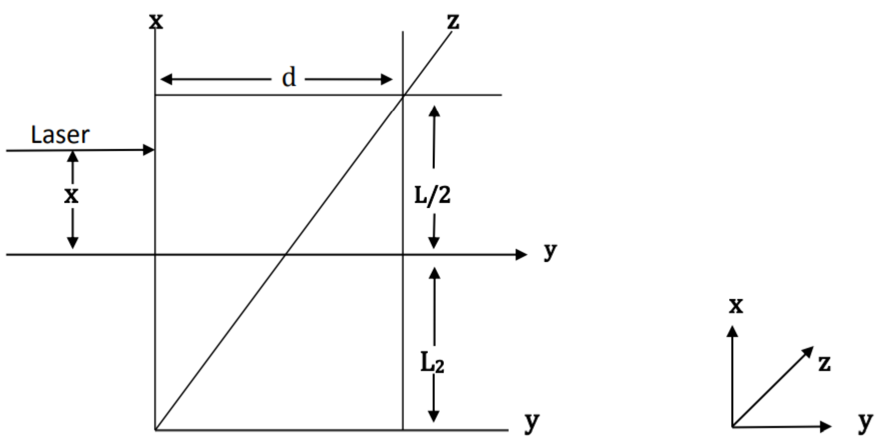
Babinet compensator is an optical device which produces desirable path difference between ordinary and extraordinary rays for any wavelengths.

6.8.2 Theory of Babinet Compensator

A Babinet compensator is a continuously variable, zero order retarder. It consists of two birefringent wedges, one of which is moveable, and another is fixed to a compensator plate. A Babinet compensator is construction from two pieces of birefringent optical material (quartz prism) with indices n_0 and n_e for light polarized perpendicular and parallel to the optic axis respectively .This device can be inclined towards positive value or negative value as per adjustment. Half wave plate or quarter wave plate is place in device for wavelength. A narrow laser beam with wavelength of ƛ is linearly polarized in the X Z plane at 450 to X and propagates through the compensator from left to right along the Y-axis.

(i) For d << 1 , calculate the relative phase shift of the X and Y polarized components of the exit beam in terms of n_0 , n_e , ƛ , 1, d , x .

(ii) Find the value of x, for the case that the emerging light is linearly polarized and circularly polarized. If the path difference in integral multiple of wavelength Δ = nƛ .The path difference due to second birefringent material is Δ = (ƛ/ß)dß



Formula used $(n_0\text{-}n_e) = \lambda d\beta/\beta t$

- Where $d\beta$ is the fringe shift with second material
- β is fringe width without the second material
- t is the thickness of the second material
- λ is wavelength of the light used
- $(n_0\text{-}n_e)$ is difference in the refractive indices of the O and E rays.

Working:- It is a constantly varying wave plate of zero order. It is used for spectrum range to achieve retardance. Long birefringes wedges position for attaining desired retardance. Now, set up the apparatus then using the micrometer of the Babinet Compensator, the fringe width is measured.A mica sheet is introduced between the polarizer and compensator so there will be fringe shift. Measure the fringe shift through micrometer for different fringes and time the fringe shift is calculated the birefringence by using the above formula.

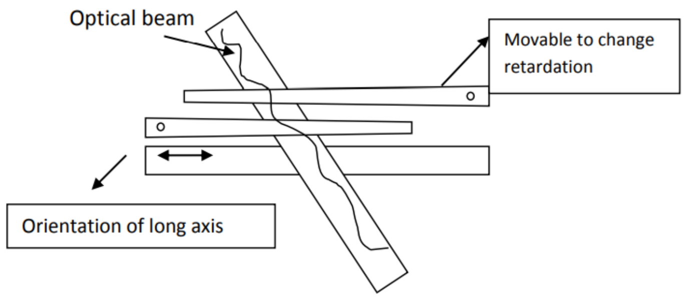


Figure : 9

6.8.3 Application of Babinet Compensator

Reorientation of the optical elements in a standard experimental setup that utilizes a Babinet compensator **to measure the anomalous dispersion of a birefringent mediumn** results in a useful trade-off between the intensity and visibility of the polarization fringes produced by the apparatus.

6.9 WHAT IS OPTICAL ROTATION

The phenomenon of optical rotation was studied in detail by Biot in the year 1815, and he proposed the laws corresponding to optical rotation. The laws of optical rotations are given as follows:

1. The amount of optical rotation produced by optically active crystals or substances is directly proportional to the thickness of the crystal or the path length traversed in its rotation.

   $\theta \propto L$

   Where,

   θ - The angle of optical rotation

   L - The path length

2. The amount of optical rotation is directly proportional to the concentration of the optically active solution.

   $\theta \propto C$

   Where,

   C - Concentration of the optically active solution

3. The angle of optical rotation is inversely proportional to the square of the wavelength of the light used.

$$\theta \; \alpha \, \frac{1}{\lambda^2}$$

Where,

λ - The wavelength of the light used

Combining all these points we find that the angle of rotation or formula for rotation is given by :

$\theta \; \propto CL$

Where,

θ - The angle of optical rotation

L – The path length

C - Concentration of the optically active solution

Further, the proportionality constant is replaced by a constant S known as the specific rotation. Then the formula for optical rotation or angle of optical rotation is given by:

θ = SCL

Where,

S-Specific rotation

Note: The angle of rotation is also depending upon the temperature and nature of the optically active substances.

6.10 WHAT IS SPECIFIC ROTATION

The angle of rotation produced by an optically active substance depends on

(1) thickness of a solid substance or length of a solution,

(2) density of a solid substance or concentration of a solution,

(3) temperature

(4) wavelength of light used.

For an optically active solid at a given temperature and for a given wavelength the angle of rotation produced (θ) is directly proportional to the thickness (t)

That is $\theta \propto t$ or $\theta = S\,t$

The constant S is called the specific rotation or specific rotatory power of the substance. It is given by $S = \frac{\theta}{t}$. Its unit is $rad\,m^{-1}$.

Specific rotation of an optically active solid substance is defined as the angle of rotation of the plane of polarisation produced by the substance of unit thickness a given temperature and for a given wavelength of light.

In case of an optically active solution, the angle of rotation produced (θ) is directly proportional to (1) the length of the solution (L) and (2) concentration of the solution (C), provided the temperature and wavelength of light remain the same.

That is $\theta \propto LC$ or $\theta = SLC$ .The constant S is called the specific rotation or specific rotatory power of the solution and is given by $S = \frac{\theta}{LC}$. Its unit is $rad\ m^2\ kg^{-1}$.

The specific rotatory power of an optically active solution is defined as the angle of rotation of the plane of polarisation produced by the solution of unit length and unit concentration at a given temperature and for a given wavelength of light

6.11 OPTICAL ACTIVITY

The property of a substance by virtue of which it rotates the plane of polarization the light incident on it is called **optical activity**. The substances having this property are called optically active substances. Eg. Quartz, sugar solution, sodium chlorate, quinine ete…

There are two types of optically active substances.

1. **Dextro-rotatory** or right handed substances are those which rotate the plane of polarisation in the clockwise direction as seen from the emergent side. Eg. Cane sugar.

2. **Laevo-rotatory o**r left handed substances are those which rotate the plane of polarisation in the anticlockwise direction as seen from the emergent side. Eg. Fruit sugar.

6.12 FRESNAL THEORY OF OPTICAL ROTATON

1. Fresnel theory of optical activity is based on the fact that when plane polarized light is allowed to pass through a crystal along the optic axis, it is split into two circularly polarized vibrations rotating in opposite directions with the same frequency and each with an amplitude half that of the incident light.
2. The velocities of the component wavers are the same in an optically inactive crystal whereas the velocities are different in an optically active crystal.
3. Calcite crystal is not optically active.

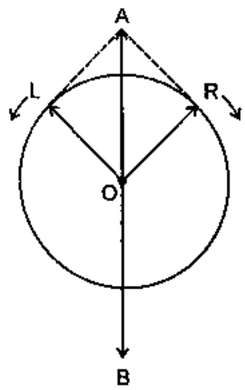


Figure : 10

4. The two component circularly polarized vibrations move forward with same velocity and therefore have a phase of /2 .
5. Let OL be the component of circularly polarized vector rotating to the left (anticlockwise) and OR the component rotaing to the right (clockwise). The resultant will be along OA. Thus there is no rotation.
6. Quartz crystals are optically active. The two component circularly polarized vibrations move forward with slightly different velocities and therefore have definite phase difference between them.
7. In case of a right handed quartz crystal, the velocity of clockwise rotation OR is more than the anti-clockwise rotation OL. Thus in a

dextro-rotatory quartz crystal, the resultant of OR and OL will be along OA'. The incident plane polarized vibrations are along AB. Thus the plane of vibration of light is rotated through an angle $\phi/2$ on passing through the crystal as shown. This angle of rotation depends on the thickness of the crystal.

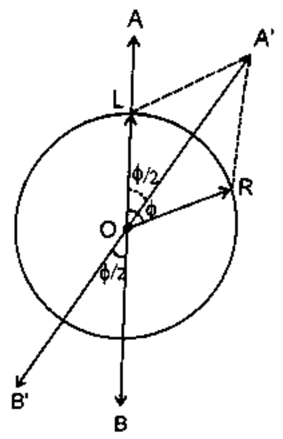


Figure : 11

6.13 WHAT IS POLARIMETERS

- Polarimetry is the measurement and interpretation of the polarization of transverse waves.
- Polarimetry is one of the important instrumental method employed in the analysis.
- Polarimetry is a sensitive, non-destructive technique for measuring the optical activity compounds.
- This technique involves the measurement of change in the direction of vibration of polarized light when interact with an optically active compound.
- A substance is said to be optically active if it rotates the plane of the polarized light.
- Instrument measures the rotation of polarized light as it passes through an optically active substance and the tendency of the molecule to rotate the plane

polarized light towards clock-wise oranti-clock wise direction whose extent of the rotation can be measured.

- In principle, a pair of crossed polarizers (a pair with their passaxes perpendicular to each other) may be used as polarimeter. Polarimeter : axes perpendicular to each other) may be used as polarimeter.
- No light will emerge from such a combination.
- If an optically active substance is introduced between them, the plane of polarization of the light emerging from it may be rotated by a certain angle (let it be) and the second polarizer will not be able to block the light now. α
- The second polarizer will have to be rotated by an angle in the same sense to make the field of view dark again. The angle of rotation can thus be measured by fitting a circular scale to the second polarizer.

6.13.1 Half Shade Polarimeters

- Consists of a semi-circular half wave plate ABC of quartz (cut parallel to optic axis) so that it introduces a phase change of π between the extra-ordinary and the ordinary rays passing through it, and a semi-circular glass plate ADC.

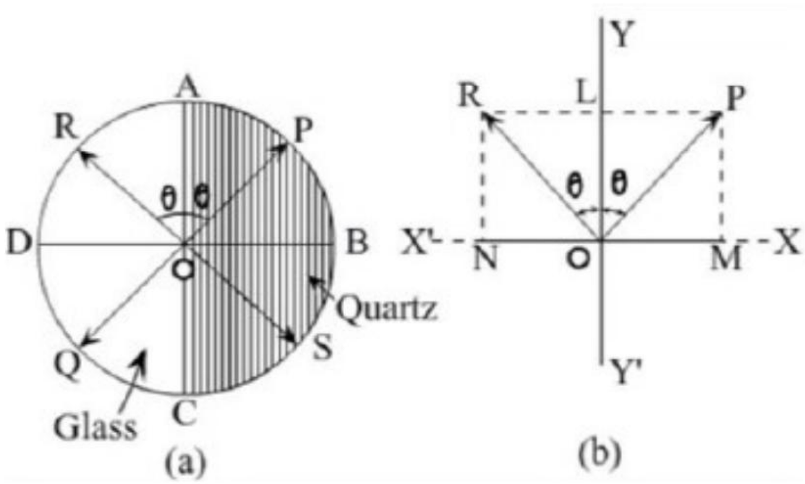


Figure : 12

- The thickness of the glass plate is such that it absorbs same amount of light as the quartz plate.

- Let the plane a vibration of the plane polarized light incident normally on the half shade device be along PQ making an angle θ with AC. The vibrations emerge from the glass plate part of the half shade device as such i.e., there is no change along the plane PQ.

- Inside the quartz plate which is doubly refracting, the light is divided into two components as we know, one ordinary component XX' and the other extraordinary component parallel to the optic axis along YY' .

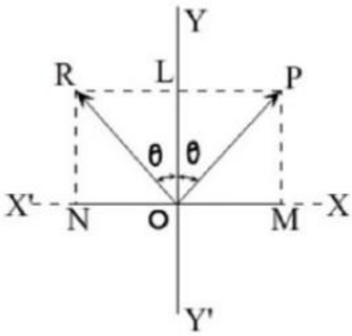


Figure : 13

- The two components travel along the same direction through separation but with different velocities. The ordinary component moves with greater velocity than the extraordinary component.

- On emergence, a phase difference of π is introduced between them.

- Due to this phase difference the direction of the ordinary component gets reversed. If the initial position of the ordinary component is represented by OM, then the final position is represented by ON. Now, on emergence the resultant of the extraordinary OL and ordinary component ON will be OR making an angle theta with the y axis.

- The vibrations of the beam emerging out of the quartz portion of the half shade device will be along RS. That is the change.

- The essential parts of this polarimeter are a monochromatic light source, a convex lens which changes the incident light beam into a parallel one. A polarization which makes this light plane polarized then, the Laurent's half shade device and then a tube containing the optically active experimental substance.

- The light beam emerging from this cube passes through an analyzer. This analyzer is capable of rotation about a common axis. The rotation of the analyzer can be read on a circular scale fitted with verniers. The light after passing through the analyzer is viewed through a telescope which is focused on the half shade device. If the pass direction of the analyzing polarizer which is capable of being rotated and if it is fitted with the circular scale.

- When the past direction is parallel to PQ then light from the glass portion will pass unobstructed while the light from the quartz portion will be partly obstructed. Due to this, the last half will appear brighter than the quartz half.

- On the other hand, if the pass direction of the analyzer is parallel to RS, light from the quartz portion will pass unobstructed, but the light from the glass portion will be partly obstructed. Thus the quartz half will appear brighter than the glass half. If however, the past direction of the analyzer is parallel to AC, y axis it equally inclined to the two planes polarized lights.

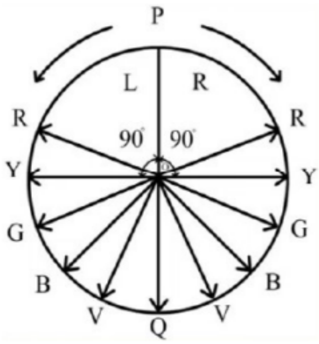
- Hence the field of view, in the two halves, will be equally bright because the half shade device serves the purpose of dividing the field of view in 2 halves. A little change in the direction of the pass direction of the analyzer, makes one half brighter other half darker.

- In the experiment to begin with, the experimental tube is filled with water; the telescope is focused on the half shade device and analyzer is rotated till the 2 halves are equally bright. This position is noted on the circular scale.

- The tube is then filled with the optically active solution and placed in position. The analyzer is rotated and it brought to a position such that the two halves are equally bright again. This new position is noted and the difference between the two regions gives the angle of rotation, pretty accurately.

6.13.2 Biquartz Shade Polarimeters

- It consists of a white light source, a convex lens as before, which renders the light into a parallel beam, the polarizer changes the incident beam into a plane polarized beam, then a biquartz plate, then the experimental tube containing the active substance and a polarizer working as an analyzer as before. A telescope is fitted with a circular scale and is focused on the biquartz plate.



- Biquartz plate consists of two semicircular plates of quartz. One of left handed quartz L and other of right handed quartz R, each of thickness about 3.75 millimeters. Both are cut perpendicular to the optic axis. This means that propagation here is along the optic axis now. They are joined together along the diameter PQ.

- When the plane polarized white light passes through a biquartz plate normally, along the optic axis the phenomena of rotary dispersion occurs because the planes of vibrations of different colors are rotated through different angles. Remember, we have seen that the amount of rotation is proportional to 1 upon lambda square and rotation will be in one sense for the left handed portion and other direction for the right handed portion. The amount of rotation is maximum for violet which has the minimum wavelength and least for red.

- The sense of rotation is opposite in the two halves. The amount of rotation also depends on the thickness. For a thickness of 3.75 millimeters, the rotation of the plane of polarization for yellow light is about 90 degrees. Hence YOY is a straight line.

- If the pass direction of the analyzer is parallel to POQ, the yellow light will not be transmitted through the analyzer, Malus Law and the appearance of two halves will be similar. The two halves will have a grayish violet tint called the tint of the passage. When the analyzer is rotated to one side from this position, one half of the field of view appears blue, while the other half appears red.

- When the analyzer is rotated in the opposite direction, the colors are interchange the first half which was bluish earlier now appears red and the second half which was reddish earlier now appears blue.

- The position dependence of the tint of passage is very sensitive and is used for accurate determination of the angle of optical rotation.

6.14 SUMMARY

Polarization is a phenomenon caused by the wave nature of electromagnetic radiation, according to physics. Sunlight is an example of an electromagnetic wave because it travels through the vacuum to reach the Earth. Because an electric field interacts with a magnetic field, these waves are known as electromagnetic waves.

Plane polarised light is made up of waves with the same direction of vibration for all of them. Light can be polarised through reflection or passing through filters, such as crystals, that transmit vibration in one plane but not in others.

The light is referred to as plane or linearly polarised with respect to the direction of propagation if the electric field vectors are restricted to a single plane by filtration of the beam with specialised materials, and all waves vibrating in a single plane are referred to as plane parallel or plane-polarized.

Humans can perceive polarisation of light, despite the fact that most of us are unaware of our ability to do so. We use 'Haidinger's brushes,' an entoptic visual phenomenon described by Wilhelm Karl von Haidinger in 1844, to detect the orientation of polarised light.There are a few different methods for polarising light.

A **waveplate** or **retarder** is an optical device that alters the polarization state of a light wave travelling through it. Two common types of waveplates are the *half-wave plate*, which shifts the polarization direction of linearly polarized light, and the *quarter-wave plate*, which converts linearly polarized light into circularly polarized light and vice versa. A quarter-wave plate can be used to produce elliptical polarization as well.

Waveplates are constructed out of a birefringent material (such as quartz or mica, or even plastic), for which the index of refraction is different for light linearly polarized along one or the other of two certain perpendicular crystal axes. The behavior of a waveplate (that is, whether it is a half-wave plate, a quarter-wave plate, etc.) depends on the thickness of the crystal, the wavelength of light, and the variation of the index of refraction. By appropriate choice of the relationship between these parameters, it is possible to introduce a controlled phase shift between the two polarization components of a light wave, thereby altering its polarization.

6.15 TERMINAL QUESTIONS

1. Explain the Principle and Construction of a quarter wave plate and half wave plate.
2. Define Specific Rotation.
3. What is Plane Polarized Light?
4. What is Babinet's Compensator? Write down the application of Babinet Compensator.
5. Explain the concept of Polarimeters.

6.16 SOLUTION & ANSWER OF TERMINAL QUESTIONS

1. Section 6.4.1, 6.4.2
2. Section 6.9
3. Section 6.3.1
4. Section 6.7, 6.7.3
5. Section 6.12

6.17 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.
2. Optics – Ajoy Ghatak
3. Optics – Principles and Application, K. K. Sharma, Academic Press.
4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.
5. Advanced Geometrical Optics – Psang Dain Lin.

DCEPHS-105

OPTICS

Uttar Pradesh Rajarshi Tandon

Open University

Bachelor of Science

Block

3 Advanced Digital Electronics

Course Design Committee

Prof. Ashutosh Gupta	**Chairman**
Director, School of Science, UPRTOU, Prayagraj	
Prof. A. K. Rai	**Member**
Professor, Dept. of Physics, University of Allahabad, Prayagraj	
Gorakhpur	
Prof. Ram Kripal	**Member**
Professor, Dept. of Physics, University of Allahabad, Prayagraj	
Dr. Anjani Kumar Singh	**Member**
Associate Professor (Retd.), E.C.C University of Allahabad, Prayagraj	

Course Preparation Committee

Dr. Akhilesh Kumar Singh	**Author**
Associate Prof.	**Unit- 02-03**
Dept. of Physics	
Pt. DDUGG, PG college Rajaji Puram, Lucknow	
Dr. Ratana Kumar	**Author**
Asst. Prof.	**Unit- 06-10**
Dept. of Physics	
RIET,college , Phagwara , Punjab	
Dr. Arvind Kumar Mishra	**Author**
Assistant Professor , Physics,	**Unit- 01, 04 & 05**
School of Science, UPRTOU, Prayagraj	
Prof. Nitya Kaushal Kishore	**Editor**
Ex. HOD, Dept. of Physics,	**Block -I,II & III**
S.K.M.U. University,	**Unit- 01 to 10**
Dumka, Jharkhand.	
Dr. Arvind Kumar Mishra	**Coordinator**
Assistant Professor , Physics,	
School of Science, UPRTOU, Prayagraj	

DCEPHS – 105 : Optics
ISBN-

Printed By: Chandrakala Universal Pvt. 42/7 Jawahar Lal Nehru Road, Prayagraj.

UNIT 07: CONCEPT OF INTERFERENCE

STRUCTURE:

7.1 INTRODUCTION

In the present unit we will discuss about concept of interference, as well know that the phenomenon of interference of light has proved the validity of the wave theory of light. Thomas Young successfully demonstrated his experiment on interference of light in 1802. When two or more wave trains act simultaneously on any particle in a medium, the displacement of the particle at any instant is due to the superposition of all the wave trains. Also, after the superposition, at the region of cross over, the wave trains emerge as if they have not interfered at all. Each wave train retains its individual characteristics. Each wave train behaves as if others are absent. This principle as explained by Huygegns in 1678.

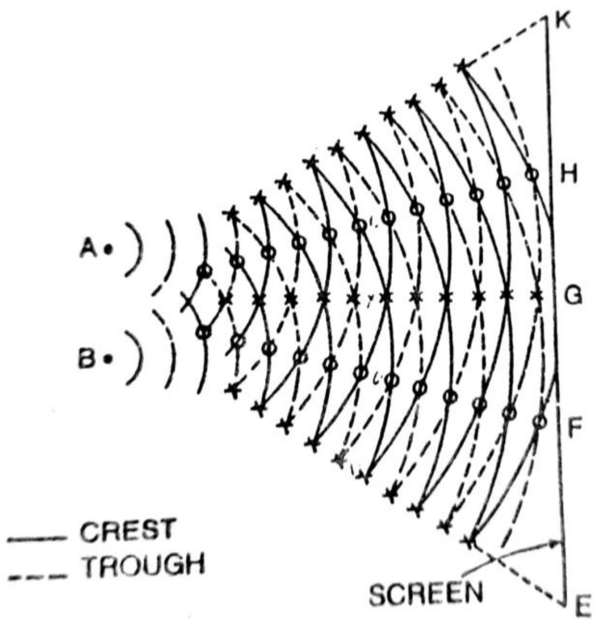


Figure : 1

The phenomenon of interference of light is due to the superposition of two trains within the region of cross over. Let us consider the waves produced on the surface of water. In Fig.1 points A and B are the two sources which produce wave of equal amplitude and constant phase difference. Waves spread out on the surface of water which are circular in shape. At any instant, the particle will be under the action of the displacement due to both the waves. The points shown by circles in the diagram will have minimum displacement because the crest of one wave falls on the trough of the other and the resultant displacement is zero. The points shown by crosses in the diagram will have maximum displacement because, either the crest of one will combine with the crest of the other or the trough of one will combine with the trough of the other. In such a case, the amplitude of the displacement is twice the amplitude of either of the waves. Therefore, at these points the waves reinforce with each other. As the intensity (energy) is directly proportional to the square of the amplitude $(I \propto A^2)$ the intensity at these points is four times the intensity due to one wave. It should be remembered that there is no loss of energy due to interference. The energy is only transferred from the points of minimum displacement to the points of maximum displacement.

7.2 OBJECTIVES

After studying this unit, student will be able to understand :

- Statement of Principle of Superposition.
- Concept of Interference.
- Fringe in Wave optics explanation.
- Young's double slit experiment and explanation.
- Explain the concept of Fresnel Biprism.
- Discuss Lloyd's Single Mirror and Fresnel's Double Mirror.
- Different type of Fringes.

7.3 PRINCIPLE OF SUPERPOSITION

According to the principle of superposition. The resultant displacement of a number of waves in a medium at a particular point is the vector sum of the individual displacements produced by each of the waves at that point.

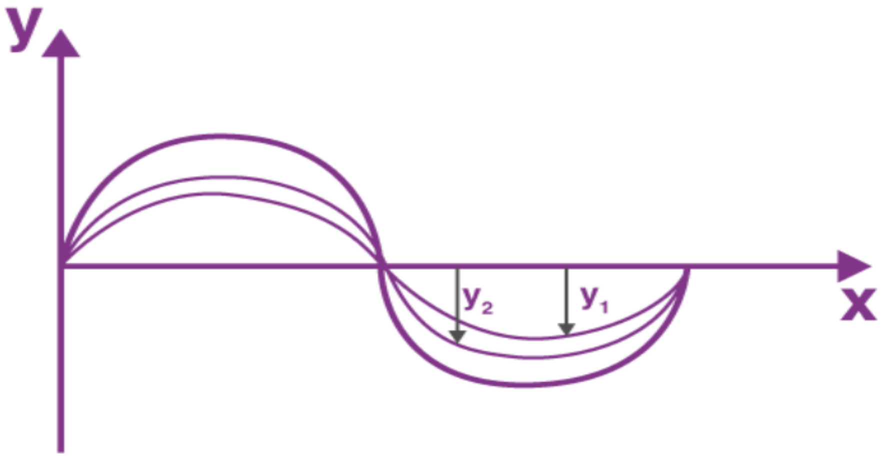


Figure : 2

And according to this, the net displacement of any component on the string for a given time is equal to the algebraic totality of the displacements caused due to each wave. Hence, this method of adding up individual waveforms for the evaluation of net waveform is termed as the **principle of superposition**.

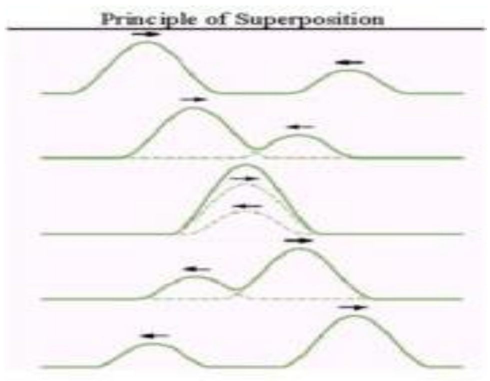


Figure : 3

The principle of superposition is expressed by affirming that overlapping waves add algebraically to create a resultant wave. Based on the principle, the overlapping waves (with the frequencies $f_1, f_2 \ldots, f_n$) do not hamper the motion or travel of each other. Therefore, the wave function (y) labeling the disturbance in the medium can be denoted as:

$$y = f_1(x - vt) + f_2(x - vt) + \ \ldots + f_n(x - vt)$$

$$= \sum_{i=1}^{n} f_i\,(x - vt)$$

Hence, the superposition of waves can lead to the following three effects:

1. Whenever two waves having the same frequency travel with the same speed along the same direction in a specific medium, then they superpose and create an effect known as the interference of waves.
2. In a situation where two waves having similar frequencies move with the same speed along opposite directions in a specific medium, then they superpose to produce stationary waves.
3. Finally, when two waves having slightly varying frequencies travel with the same speed along the same direction in a specific medium, they superpose to produce beats.

Concept of Principle of Superposition

Suppose there are two sources of waves S_1 and S_2.

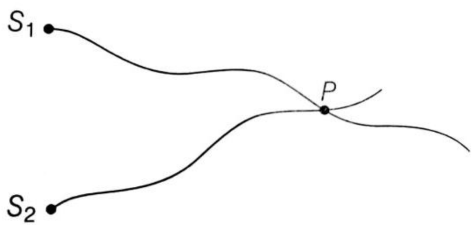


Figure : 4

Now, the two waves from S1 and S2 meet at some point (say P). Then, according to principle of superposition net displacement at P (from its mean position) at any time is given by

$$\boxed{y = y_1 + y_2}$$

Here, y_1 and y_2 are the displacement of P due to two waves individually.

For example, suppose at 9 AM, displacement of P above its mean position should be 6 mm accordingly to wave-1 and at the same time its displacement should be 2 mm below its mean position accordingly to wave-2, then at 9 AM net displacement of P will be 4 mm above its mean position.

Now, based upon the principle of superposition we have two phenomena in physics, interference and beats. Stationary waves (or standing waves) and Young's double slit experiment (or YDSE) are two examples of interference.

Based on principle of superposition means two or more than two waves meet at one point or several points and at every point net displacement is $y = y_1 + y_2$ or $y = y_1 + y_2 + y_3$ etc.

7.3.1 Resultant Amplitude and Intensity due to Coherent Sources

We have seen that the two waves from two sources S1 and S2 were meeting at point P. Suppose they meet at P in a phase difference $\Delta\phi$ (or ϕ). If this phase different remains constant with time, then sources are called coherent, otherwise incoherent.

For sources to be coherent, the frequencies (f, ω or T) of the two sources must be same. This can be understood by the following example.

Suppose the phase difference is 0^o. It means they are in same phase. Both reaches their extremes ($+ A$ or $-A$) simultaneously. They cross their mean positions (in the same direction) simultaneously. Now, if we want their phase difference to remain constant or we want that above situation is maintained all the time, then obviously their time periods (or frequencies) must be same.

Resultant Amplitude

1. Consider the superposition of two sinusoidal waves of same frequency (means sources are coherent) at some point. Let us assume that the two waves are travelling in the same direction with same velocity. The equation of the two waves reaching at a point can be written as

$$y_1 = A_1 \sin(kx - \omega t)$$

and $$y_2 = A_2 \sin(kx - \omega t + \phi)$$

The resultant displacement of the point where the waves meet, is

$y = y_1 + y_2$

$= A_1 \sin(kx - \omega t) + A_2 \sin(kx - \omega t + \phi)$

$$= A_1 \sin(kx - \omega t) + A_2 \sin(kx - \omega t) \cos\phi + A_2 \cos(kx - \omega t) \sin\phi$$

$$= (A_1 + A_2 \cos\phi) \sin(kx - \omega t) + \sin\phi \cos(kx - \omega t)$$

$$= A \cos\theta \sin(kx - \omega t) + A \sin\theta \, cos(kx - \omega t)$$

or $y = A\, sin\, (kx - \omega t + \theta)$

Here, $A_1 + A_2 cos\, \phi = A\, cos\theta$

and $A_2 \sin\phi = A \sin\theta$

or $A^2 = (A_1 + A_2 \cos\phi)^2 + (A_2\, sin\phi)^2$

or $\boxed{A = \sqrt{A_1^2 + A_2^2 + 2A_1A_2\, cos\, \phi}}$ …………

(1)

and $\tan\theta = \frac{A\sin\theta}{A\cos\theta} = \frac{A_2\sin\phi}{A_1 + A_2\cos\phi}$

2. The above result can be obtained by graphical method also. Assume a vector A_1 of length A_1 to represent the amplitude of first wave.

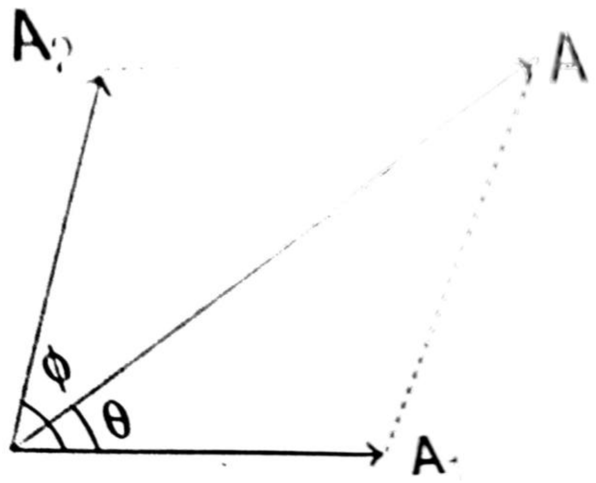


Figure : 5

Another vector A_2 of length A_2, making an angle ϕ with A_1 represent the amplitude of second wave. The resultant of A1 and A2 represents the amplitude of resulting function y. The angle θ represents the phase differences between the resulting function and the first wave.

Resultant Intensity

In the previous chapter, we have read that intensity of a wave is given by

$$I = \frac{1}{2}\rho\omega^2 A^2 v \quad or\ I \propto A^2$$

So, if ρ, ω and v are same for the both interfering waves then Eq. (1) can also written as

$$\boxed{I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\phi} \qquad \text{………..}$$

(2)

Here, proportionality constant $(I \propto A^2)$ cancels out on right hand side and left hand side.

Note: (i) Eqs. (1) and (2) are two equations for finding resultant amplitude and resultant intensity at some point due to two coherent sources.

(ii) In the above equations ϕ is the constant phase difference at that point as the sources are coherent. Value of this constant phase difference will be different at different points.

(iii) The special case of above two equations is, when the individual amplitudes (or intensities) are equal.

(iv) The special case of above two equations is, when the individual (or intensities) are equal or $A_1 = A_2 = A_0\ (say)$ …….. (3)

∴ $l_1 = l_2 = l_0\ (say)$ ……..

(4)

In this case, Eqs. (1) and (2) become

And
$$\boxed{\begin{gathered} A = 2A_0\ cos\frac{\phi}{2} \\ l = l_0 \cos\frac{\phi}{2} \end{gathered}}$$

(v) From Eqs. (1) to (4) we can see that, for given values of A_1, A_2, l_1 and l_2 the resultant amplitude and the resultant intensity are the functions of only ϕ.

(vi) If three or more than three waves (due to coherent sources) meet at some point then there is no direct formula for finding resultant amplitude intensity. In this case, first of all we will find resultant amplitude by vector method (either by using polygon law of vector of vector addition or component method) and then by the relation $l \propto A^2$, we can also determine the resultant intensity.

For example, if resultant amplitude comes out to be $\sqrt{2}$ times then resultant intensity will become two times.

7.3.2 What is Interference

Wave interference is the phenomenon that occurs when two waves meet while traveling along the same medium. The interference of waves causes the medium to take on a shape that results from the net effect of the two individual waves upon the particles of the medium. To begin our exploration of wave interference, consider two pulses of the same amplitude traveling in different directions along the same medium. Let's suppose that each displaced upward 1 unit at its crest and has the shape of a sine wave. As the sine pulses move towards each other, there will eventually be a moment in time when they are completely overlapped. At that moment, the resulting shape of the medium would be an upward displaced sine pulse with an amplitude of 2 units. The diagrams below depict the before and during interference snapshots of the medium for two such pulses. The individual sine pulses are drawn in red and blue and the resulting displacement of the medium is drawn in green.

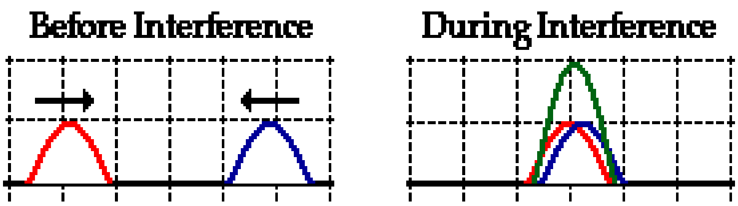


Figure : 6

Concept of Interference

For interference phenomena to take place, sources must be coherent. So, phase difference at some point should remain constant. Value of this constant phase difference will be different at different points. And since the sources are coherent, therefore following four equations can be applied for finding resultant amplitude and intensity (in case of two sources)

$A = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos\phi}$ ………… (5)

$I = I_1 + I_2 + 2\sqrt{I_1 I_2} cos\phi$ ………..

(6)

$A = 2A_0 \, cos\frac{\phi}{0}$ (if $A_1 = A_2 = A_0$) ……….

(7)

$A = 4I_0 \, cos^2\frac{\phi}{2}$ (if $I_1 = I_2 = I_0$) ……….

(8)

For given values of A_1, A_2, I_1 and I_2 the resultant amplitude and resultant intensity are the functions of only ϕ.

Now, suppose S_1 and S_2 are two coherent sources, then we can see that the two waves are meeting at several points (P_1, P_2, P_3 … etc.). At different points difference Δx will be different and therefore phase difference Δϕ or ϕ will also be different. Because the phase difference depends on the path difference $\left(\Delta\phi \ or \ \phi = \frac{2\pi}{\lambda}.\Delta x\right)$.

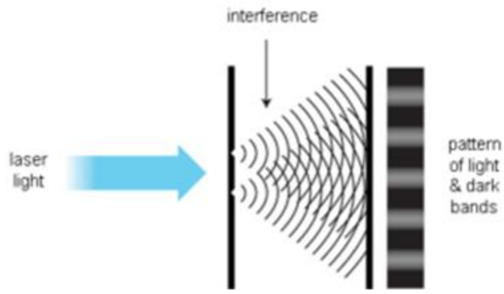


Figure : 7

Figure : 8

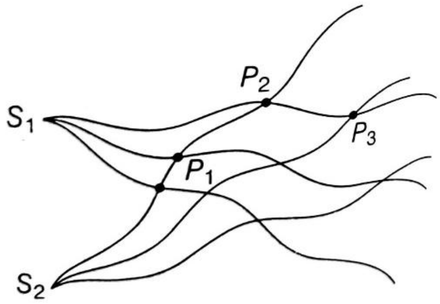


Figure : 9

And since phase difference points is different, therefore from the above four equations we can see that resultants amplitude and intensity will also be different. But whatever is the intensity at some point, it will remain constant at that point because the sources are coherent and the phase difference is constant at that point.

When Waves Meet

When two or more waves meet, they interact with each other. The interaction of waves with other waves is called **wave interference**. Wave interference may occur when two waves that are traveling in opposite directions meet. The two waves pass through each other, and this affects their amplitude. Amplitude is the maximum distance the particles of the medium move from their resting positions when a wave passes through. How amplitude is affected by wave interference depends on the type of interference. Interference can be constructive or destructive.

7.3.3 Type of Interference

The interference of light waves can be divided into the following two categories.

- **Constructive Interference:** In case of Constructive interference the crest of one wave falls on the crest of another wave in such a way that the amplitude becomes maximum. These waves are in phase with each other and have the same displacement too.

- **Destructive Interference:** In case of destructive interference, the crest of one wave falls on the trough of another wave in such a way that the amplitude becomes minimum. These waves are out of phase with each other and have different displacements.

7.3.4 Constructive Interference

Constructive interference occurs when the crests, or highest points, of one wave overlap the crests of the other wave. You can see this in the **Figure** below. As the waves pass through each other, the crests combine to produce a wave with greater amplitude.

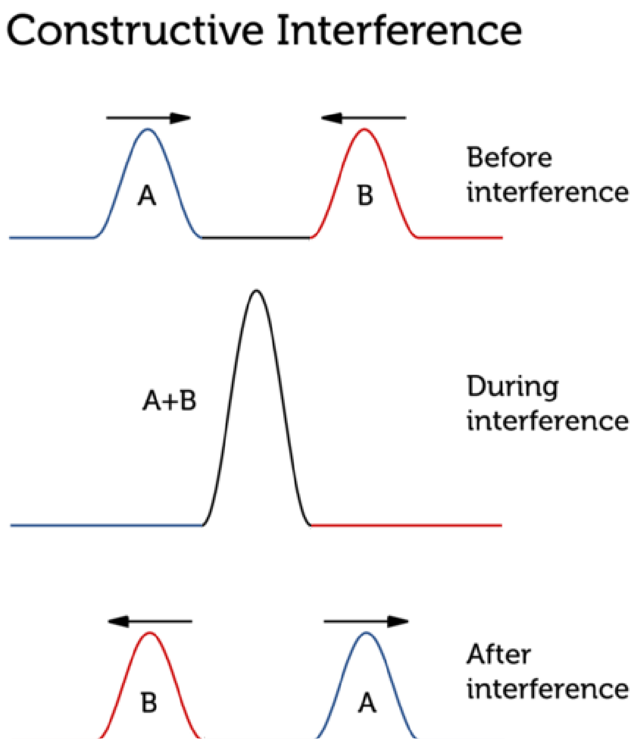


Figure : 10

7.3.5 Destructive Interference

Destructive interference occurs when the crests of one wave overlap the troughs, or lowest points, of another wave. The **Figure 11** below shows what happens. As the waves pass through each other, the crests and troughs cancel each other out to produce a wave with zero amplitude.

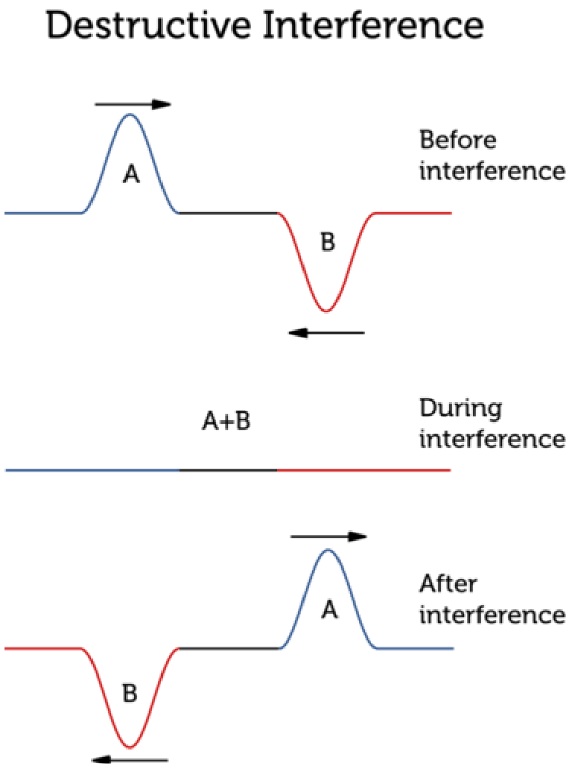


Figure : 11 (a)

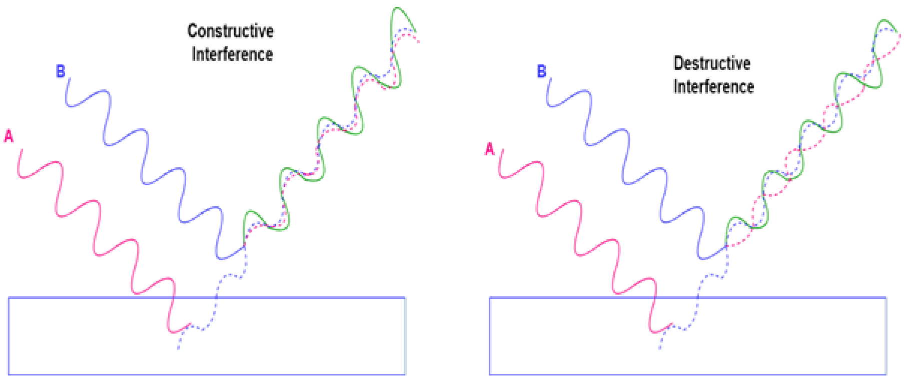


Figure 11 (b) : Constructive and Destructive Interference

7.3.6 Condition for Interference

The conditions for the interference of light are as under.

(i) In interference the source of light should be monochromatic.

(ii) Here waves should be of the same frequency.

(iii) Direction of waves should also be the same.

(iv) The amplitudes of both the waves should also be the same.

(v) The slits of both the sources should be thin. Diffraction applies to a number of phenomena that arise when a wave meets an obstacle. In classical physics, the diffraction phenomenon is defined as the apparent bending of waves around small obstacles and the spreading of waves through small openings. Different effects arise as light waves pass through a medium with different refractive indexes or a sound wave through a medium with differing acoustic impedance. Diffraction happens in all waves, including sound waves, water waves, and electromagnetic waves, such as visible light, x-ray, and radio waves. Because physical objects have wave-like properties (at the atomic level), diffraction often happens with matter and can be studied according to the principles of quantum mechanics. The Italian scientist Francesco Maria Grimaldi coined the term diffraction and was the first to report detailed observations of this phenomenon in 1665.

Remember:

Throughout physics, interference is a phenomenon in which two waves are superimposed to create a resulting wave of greater or lesser amplitude. Interference typically refers to the interaction of waves that are associated or compatible with each other, either because they come from the same source or because they have the same (or nearly the same) frequency. Interference effects can be found in all types of waves, including light, radio, acoustics and surface waves. For chemistry, the applications of light interference are the most important to the study of matter

7.4 WHAT IS FRINGE IN WAVE OPTICS?

The alternating bright and the dark band formed due to interference is called fringe. When two light waves superimpose it forms constructive interference and destructive interference. The bright band is due to constructive interference and the dark band is due to destructive interference. The width of the fringe is found using the formula

$\beta = \lambda D/d$

β is the bandwidth

λ is the wavelength of the light

D is the distance between the source and the screen

d is the distance between the two slits

7.4.1 Fringe Pattern

The alternate dark and red bands which are obtained on the screen are known as fringe pattern and the alternate dark and bright bands are known as fringes.

7.4.2 Bright Bands

Bright bands are formed as a result of constructive interference and they are the positions of maximum intensity.

Condition for maximum intensity :

Path difference $= n\lambda$

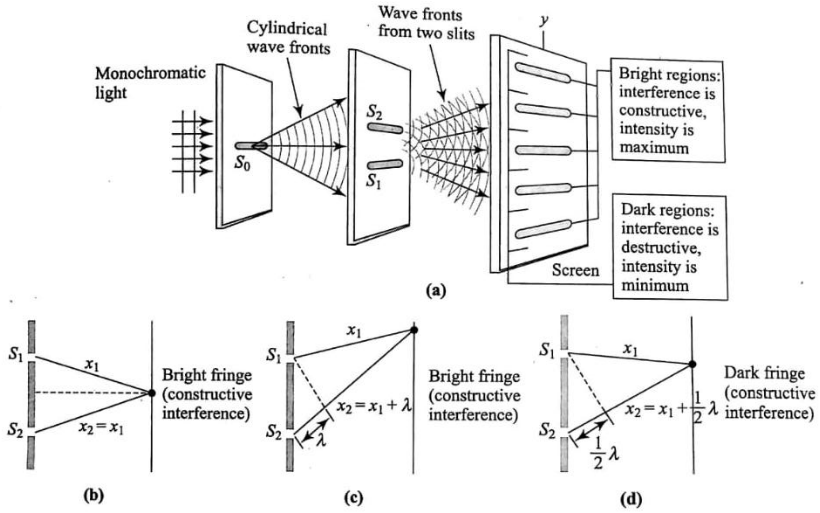
7.4.3 Dark Bands

Dark bands are formed by the destructive interference and they are the positions of minimum intensity.

7.5 YOUNG'S DOUBLE SLIT EXPERIMENT

In 1801, the English scientist Thomas Young (1773-1829) performed a historic experiment that demonstrated the wave nature of light by showing that two overlapping light waves interfered with each other. His experiment was particularly important because he was also able to determine the wavelength of the light from his measurements, the first such determination of this important property. Figure 12 (a) shows one arrangement of Young's experiment, in which light of a single wavelength (monochromatic light) passes through a single narrow slit and falls on two closely spaced narrow slits S_1, and S_2. These two slits act as coherent sources of light waves that interfere constructively and destructively at different points on the screen to produce a pattern of alternating bright and dark fringes. The purpose of the single slit is to ensure that only light from one direction falls on the double slit from different points on the light source would pattern on the screen to be washed out. The slits S_1 and S_2 act as coherent sources of light waves because the light from each originates from the same primary source, namely the single slit.

To explain the origin of the bright and dark fringes, the figure presents three top views of the double slit and the screen. Figure 12 (b) illustrates how a bright fringe arises directly opposite the midpoint between the two slits. In this part of the figure, the waves (identical) from each slit travel to the midpoint on the screen. At this location, the distances x_1 and x_2 to the slits are equal, each containing the same number of wavelengths. Therefore, constructive interference results, leading to the bright fringe. Figure 12 (c) indicates that constructive interference produces another bright fringe on one side of the midpoint when the distance x_2 is larger than x_1 by exactly one wavelength. A bright fringe also occurs symmetrically on the other side of the midpoint when the distance x_1 and x_2 by one wavelength; for clarity, however, this bright fringe is not shown. Constructive interference produces additional bright fringes on both sides of the middle wherever the difference between x_1 and x_2 is an integer number of wavelengths : λ, 2λ, 3λ, etc. Figure 12 (d) shows how the first dark fringe arises. Here, the distance x_2 is larger than x_1 by exactly one-half a wavelength, so the waves interfere destructively, giving rise to the dark fringes. Destructive interference creates additional dark fringes on both sides of the center whenever the difference between x_1 and x_2 equals an odd integer number of half-wavelengths : $1\left(\frac{\lambda}{2}, 3\left(\frac{\lambda}{2}\right)\right)$, etc.

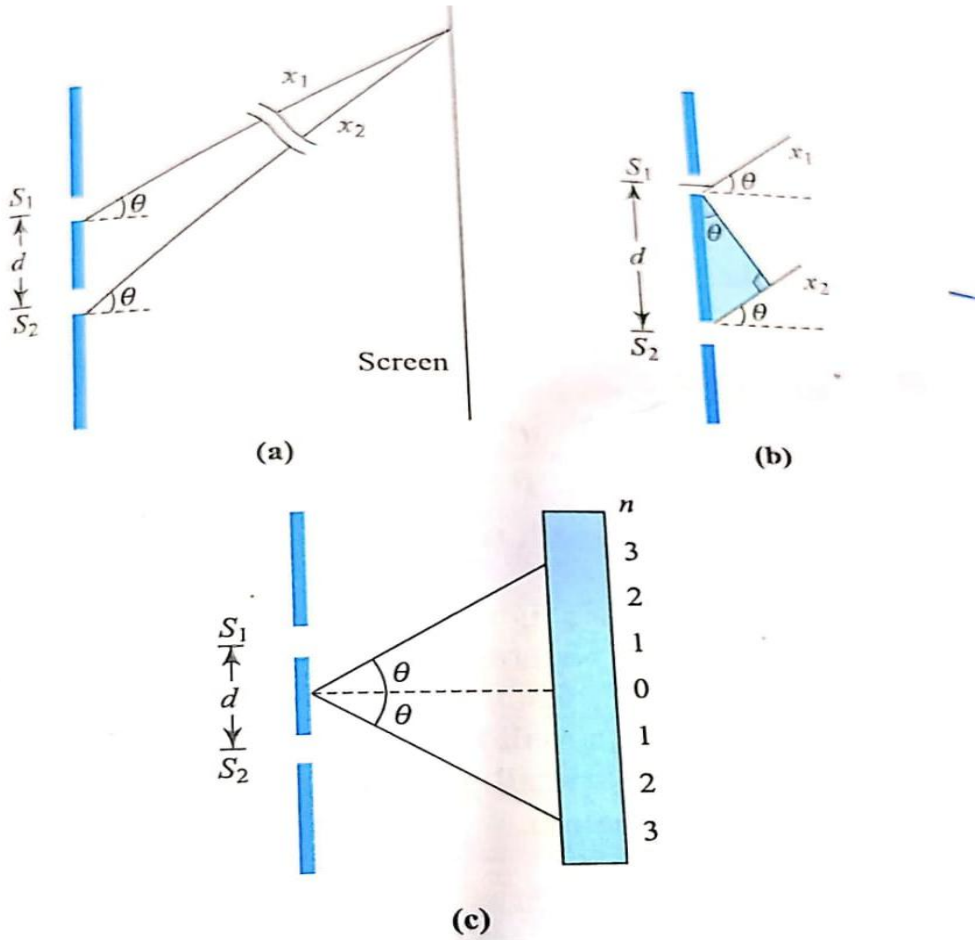


The waves slits S_1 and S_2 interfere constructively (Parts (b) and (c)) or destructively (Part (d)) on the screen, depending on the difference in distance between the slits and the screen. The slit widths and the distance between the slits have been exaggerated for clarity.

Figure : 12

The position of the fringes observed on the screen in Young's experiment can be calculated with the aid of following figure. If the screen is located far away compared with the separation d of the slits, then the lines labeled x_1 and x_2 in Fig. (a) are nearly parallel. Being nearly parallel, these lines make approximately equal angels θ with the horizontal. The distances x_1 and x_2 differ by an amount Δx, which is the length of the short side of the shaded, it follows that Δx = d sin θ. Constructive interference occurs when the distances differ by an integer number n of wavelengths λ, or Δx = d sin θ =

nλ. Therefore, the angle θ for the interference maxima can be determined from the following expression.



(a) Rays from slits S_1 and S_2, which make approxi-mately the same angle θ with the horizontal, strike a distant screen at the same spot. (b) The difference in the path lengths of the two rays is $\Delta l = d \sin\theta$. (c) The angle θ is the angle at which a bright fringe ($m = 2$, here) occurs on either side of the central bright fringe ($m = 0$).

Figure : 13

7.5.1 Bright Fringes of a Double Slit

$$\sin\theta = n\frac{\lambda}{d}, n = 0, 1, 2, 3, \ldots \qquad \ldots\ldots\ldots (9)$$

The value of n specifies the order of the fringe. Thus, n = 2 identifies the 'second-order' bright fringe. Figure 14 (a) stresses that the angle θ given by Eq. (9) locates bright fringes on either side of the midpoint between the slits. A similar line of reasoning leads to the fringes, are located according to the following expression.

7.5.2 Dark Fringes of a Double Slit

$\sin\theta = \left(n - \frac{1}{2}\right)\frac{\lambda}{d}, n = 0, 1, 2, 3, \ldots..$ ………

(10)

7.5.3 Position of Bright and Dark Fringes in YDSE

Let us consider point P on the distant screen, at a distance D from the slits width D>>d. The small are of the circle from P is almost a straight line. From figure 14, path difference $\Delta x = d \sin\theta$.

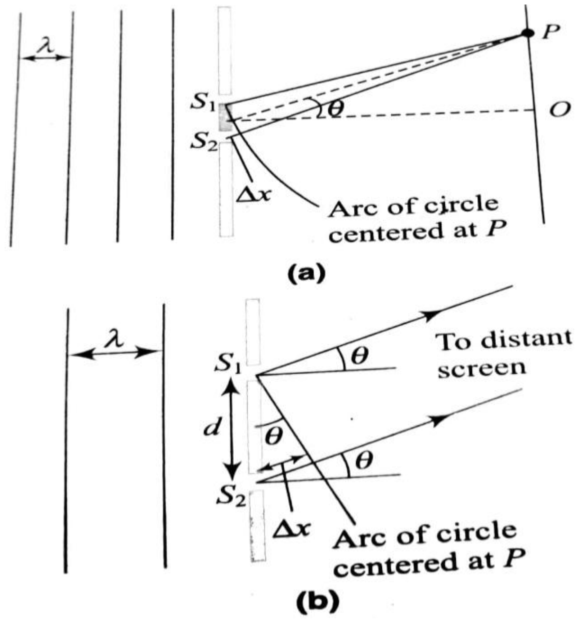


Figure : 14

The condition for constructive interference is

$$\Delta x = S_2P - S_1P = \pm n\lambda, \quad where \;\; n = 0, 1, 2, \ldots.$$

$d\sin\theta = \pm n\lambda$ (Condition for maxima) …………. (11)

The condition for destructive interference is

$$\Delta x = S_2P - S_1P = \pm\left(n - \frac{1}{2}\right)\lambda, \quad where \;\; n = 1, 2, \ldots\ldots$$

$d\sin\theta = \pm\left(n - \frac{1}{2}\right)\lambda$ (Condition for minima) …………. (12)

If the separation between screen and slits is large (D >> d), then we have

$$\sin\theta = \tan\theta = \theta = \frac{y}{D}$$

where y is the vertical distance from the center of the pattern. Position of n^{th} bright and dark fringes are, respectively,

$$\frac{y_n d}{D} = n\lambda \quad or \quad y_n = n\lambda\left(\frac{D}{d}\right)$$

and $\frac{y_n d}{D} = \left(n - \frac{1}{2}\right)\lambda \quad or \quad y_n = \left(n - \frac{1}{2}\right)\lambda\frac{D}{d}$

Each value of n corresponds to particular bright or dark fringe. The absolute value of n is called the order of interference.

The n^{th} and $(n + 1)^{th}$ maxima are given by

$y_n = \frac{n\lambda D}{d} \quad or \quad y_{n+1} = \frac{(n+1)\lambda D}{d}$ ……….

(13)

7.5.4 Fringe Width

The distance between two successive maxima (or minima) is called fringe width given as

$\beta = y_{n+1} - y_n$

Distance between $(n+1)^{th}$ maxima and n^{th} maxima will also equal to fringe width

$\beta = (n + 1)\frac{\lambda D}{d} - n\frac{\lambda D}{d}$

or $\beta = \frac{\lambda D}{d}$ ………

(14)

The fringe width is measured from centre of one fringe to centre of next fringe. The bright fringe has maximum intensity at its centre similarly dark fringes has minimum intensity at its center.

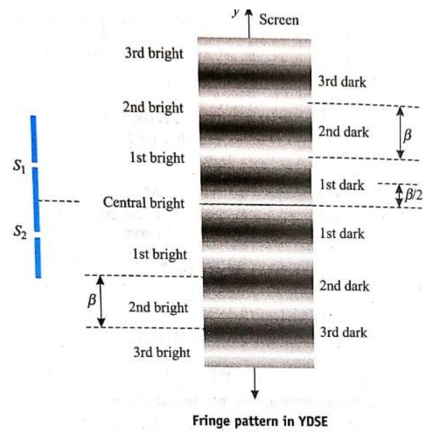


Figure : 15

7.5.5 Angular Fringe Width

The angular fringe width is $\theta = \frac{\beta}{D} = \frac{(\lambda D/d)}{D}$

or $\theta = \frac{\lambda}{d}$ …………….. (15)

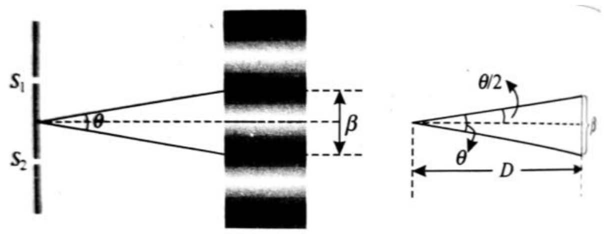


Figure : 16

The angular fringe width is defined as the ratio of wavelength of light and the distance between the slits S1 and S2 (sources) of coherent light.

7.6 WHAT IS FRESNEL BIPRISM:

Fresnel used a biprism to show interference phenomenon. The biprism consists of two prisms of very small refracting angles joined base to base. In practice, a thin glass plate is taken and one of its faces is ground and polished till a prism Fig. 17 (a) is formed with an obtuse angle of about 179^o and two side angles of the order of $30'$. When a light ray is incident on an ordinary prism, the ray is bent through an angle called the angle of deviation. As a result, the ray emerging out of the prism appears to have emanated from a prism *S'* located at a small distance above the real source, as shown in Fig. 17 (b). We say that the prism produced a virtual image of the source. A biprism, in the same way, creates two virtual sources S_1 and S_2, as seen in Fig. 17 (c). These two virtual sources are images of the same source S produced by refraction and are hence coherent.

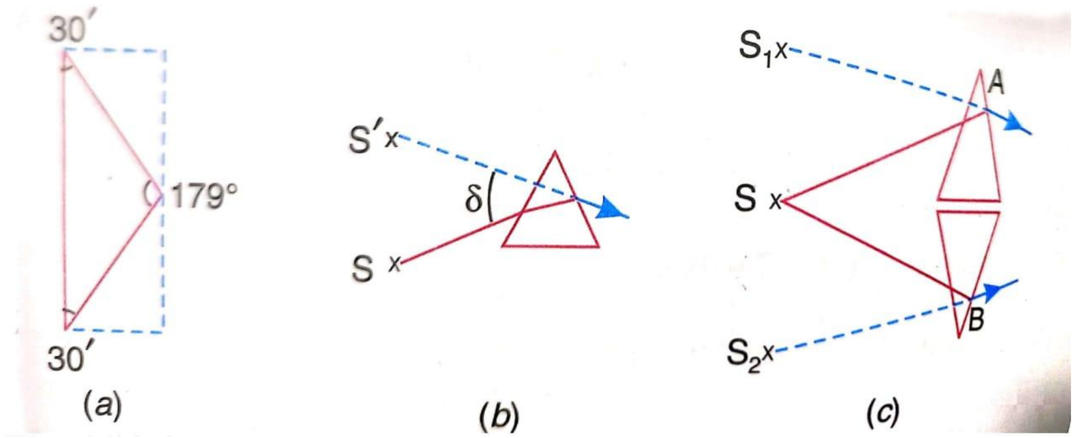


Figure : 17

7.6.1 Concept of Fresnel Biprism

The biprism is mounted suitably on an optical bench. An optical bench consists of two horizontal long rods, which are kept strictly parallel to each other and at the same level. The rods carry uprights on which the optical components are positioned. A monochromatic light source such as sodium vapors lamp illuminates a vertical slit S. Therefore, the slit S acts as a narrow linear monochromatic light source. The biprism is placed in such a way that its refracting edge is parallel to the length of the slit S. A single cylindrical wavefront impinges on both prisms. The top portion of wavefront is refracted downward and appears to have emanated from the virtual image S_1. The lower segment, falling on the lower part of the biprism, is refracted upward and appears to have emanated from the virtual source S_2. The virtual sources S1 and S2 are coherent (see Fig. 18), and hence the light waves are in a position to interfere in the region beyond the biprism. If a screen is held there, interference fringes are seen. In order to observe fringes, a micrometer eyepiece is used.

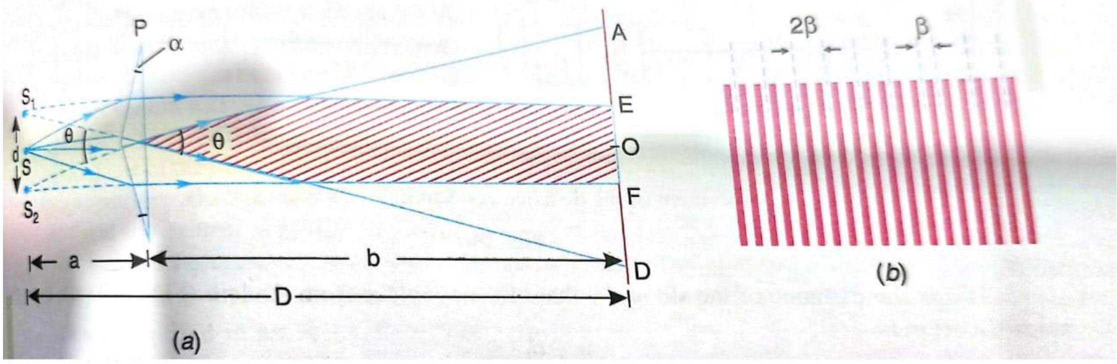


Figure : 18

Theory:

The theory of the interference and fringe formation in case of Fresnel biprism is the same as described in Section 7.5.3 for the double-slit. As the point O is equidistant from S_1 and S_2, the central bright fringe of maximum intensity occurs there. On both sides f O, alternate bright and dark fringes, as shown in Fig. 15 (b), are produced. The width of the dark or bright fringe is given by equ. 14.

$$\beta = \frac{\lambda D}{d}$$

where D (= a + b) is the distance of the sources from the eye-piece.

Figure : 19

7.6.2 Application of Fresnel's Biprism

Fresnel biprism can be used **to determine the wavelength of a light source (monochromatic), thickness of a thin transparent sheet/ thin film, refractive index of medium etc**. Biprism can be used to determine the wavelength of given monochromatic light using the expression.

7.7 LLOYD'S SINGLE MIRROR:

In 1834, Lloyd devised an interesting method of producing interference, using a single mirror and using almost grazing incidence. The Lloyd's mirror consists of a plane mirror about 30 cm in length and 6 to 8 cm in breadth (see Fig. 20). It is polished on the front surface and blackened at the back multiple reflections. A cylindrical wavefront coming from a narrow slit S_1 falls on the mirror which reflects a portion of the incident wavefront, giving rise to a virtual image of the slit S_2. Another portion of the wavefront proceeds directly from the slit S_1 to the screen. The slits S_1 and S_2 act as two coherent sources. Interference between direct and reflected waves occurs within the region of overlapping of the two beams and fringes are produced on the screen placed at a distance D from S_1 in the shaded portion EF.

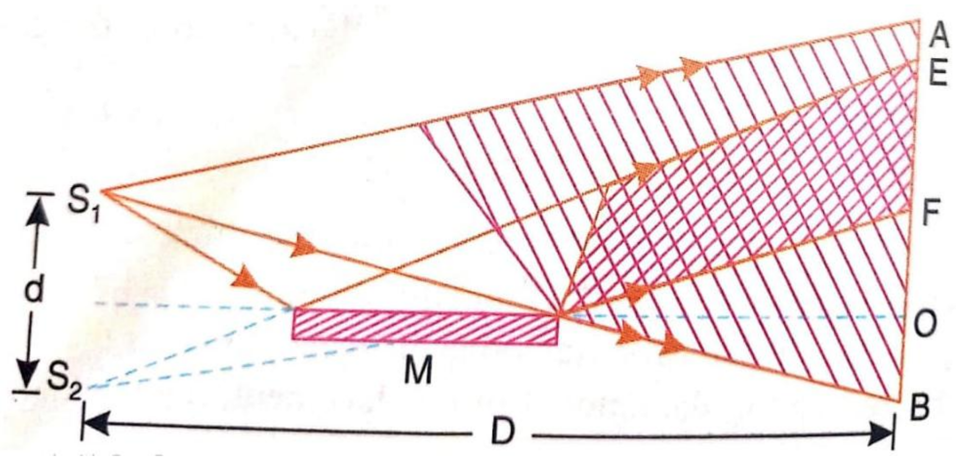


Figure : 20

The point O is equidistant from S_1 and S_2. Therefore, central (zero-order) fringe is expected to lie at O (the perpendicular dissector of S_1S_2) and it is also expected to be bright. However, it is not usually seen since the point O lies outside the region of interference (only the direct light and not the reflected light reaches O).

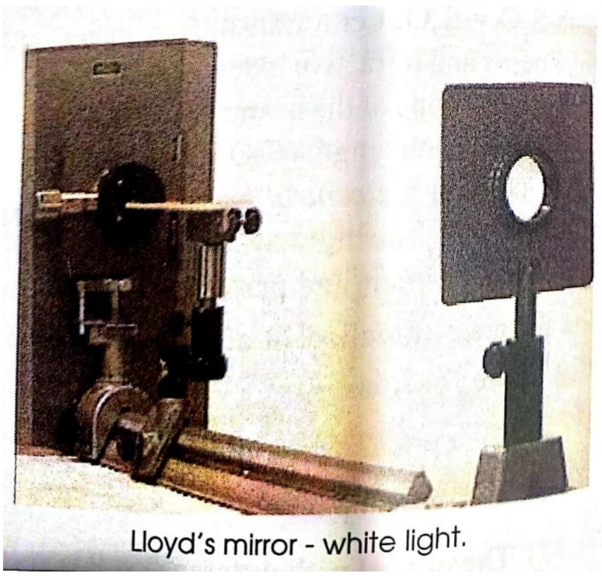


Figure : 21

By moving the screen nearer to the mirror such that it comes into contact with the mirror, the point O can be just brought into the region of interference. With white light the central fringe at O is expected to be white but in practice it is dark. The occurrence of dark fringe can be understood taking into the consideration of the phase change of π that light suffers when reflected from the mirror. The phase change leads to a path difference of λ/2 and hence destructive interference occurs there.

7.7.1 Determinations of Wavelength

The fringe width is given by equ. 14. Thus,

$$\beta = \frac{\lambda D}{d}$$

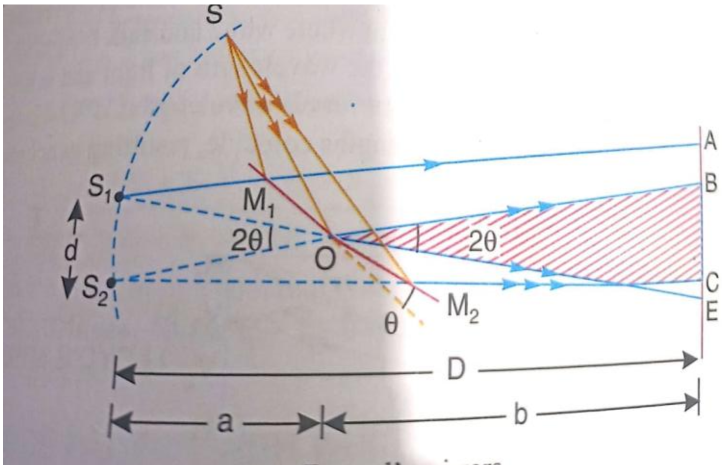
Measuring β, D and d, the wavelength λ can be determined.

7.7.2 Comparison between the fringes produced by biprism and Lloyd's mirror

1. In biprism the complete set of fringes is obtained. In Lloyd's mirror a few fringes on one side of the central fringe are observed, the central fringe being itself invisible.
2. In biprism the central fringe is bright whereas in case of Lloyd's mirror, it is dark.
3. The central fringe is less sharp in biprism than that in Lloyd's mirror.

7.8 FRESNEL'S DOUBLE MIRROR

Fresnel's double mirror is an arrangement for obtaining two coherent sources by using the phenomenon of reflection. It consists of two plane mirrors inclined to each other at a very small angle, as shown in Fig. . The mirrors are silvered on their from surfaces and are arranged at nearly 180^{o} such that their surface are nearly coplanar.



Fresnel's mirrors.

Figure : 22

A narrow slit S is placed parallel to the line of intersection of the mirror surfaces and is illuminated with monochromatic light. One portion of the cylindrical wavefront coming from slit S is reflected from the first mirror and another portion of the wavefront is reflected from the second's mirror. After reflection, the light appears to diverge from S_1 and S_2, which are the virtual images of S. As the images S_1 and S_2 of the slit are derived from the same source S, they behave as two coherent sources, placed at a distance d apart. The waves diverging from S1 and S1 overlap and interference fringes are produced in the overlapping region EF on the screen. The fringes are of equal width.

Fringe width :

It is screen from the geometry of the figure (Fig. 22) that $OS_1 = OS_2 = OS$. That is S_1, S_2 and S lie on a circle with O as a centre. Let 'a' be the distance of the sources and 'b' be the distance of the screen from O. Them the fringe width is given by

$$\beta = \frac{\lambda D}{d} = \frac{(a+b)\lambda}{d} \qquad \text{………… (16)}$$

OE and OB are the reflected rays from OM1 and OM2 respectively, corresponding to the incident ray SO. Therefore, the angle between OE and OB is twice the angle between the mirrors. Hence, $< S_1OS_2 = < BOE = 20$ Now,

$$Are\ S_1S_2 = a\ \times 2\theta$$

$$\therefore\ \ d = a\ \times 2\theta$$

Using the above result into equ. (16), we get

$$\beta = \frac{(a+b)}{2a\theta}\,\lambda \qquad \text{............ (17)}$$

Comparison between the fringes produced by biprism and double mirror :

The fringes in both cases similar in appearance. However, the double mirror fringes are narrower than the biprism fringes.

7.9 ACHROMATIC FRINGES

A system of white and dark fringes, without any colors, obtained by white light is called achromatic fringes.

When the slit is illuminated by white light in any interference experiment, we obtain a central white fringe flanked by a few coloured fringes. Coloured fringes are obtained because the fringe width is dependent on the wavelength ($\beta = \lambda D/d$). For example, the width of the red fringe is more than the blue fringe. Rayleigh designed an experiment where white and dark fringes were obtained. It can be done if the fringe width is independent of the wavelength of light and is the same for all wavelengths. The fringe width β can be kept constant for all wavelengths, if λ/d is the same in all cases. Then the maxima of each order for all wavelengths coincide, resulting in achromatic fringes.

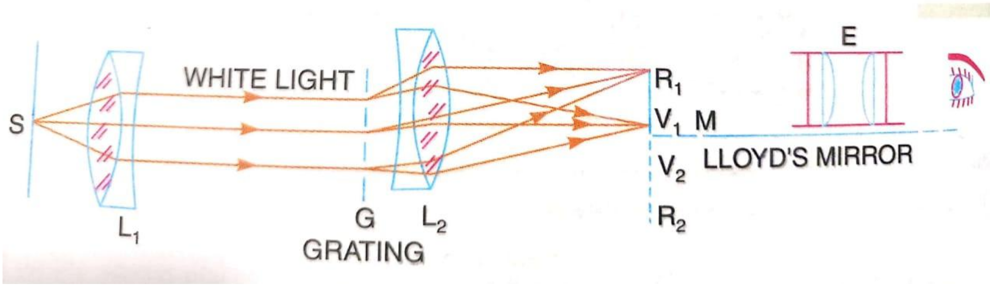


Figure : 23

In practice, achromatic fringes may be obtained as follows. S is a narrow source of white light at the focal plane of the converging lens L_1. A grating G having 800 to 1200 lines per cm is placed normal to light emerging from L_1. Another achromatic lens L_2 is used to form the second order spectrum on an opaque screen with a narrow opening in it. The narrow opening is adjusted so that only the first order spectrum is allowed to pass through it. The violet end is nearer to the highly polished Lloyd's mirror M than the red end. The position of M is adjusted such that V_2 and R_2 are the images of V_1 and R_1. Interference occurs between the beams from V_1R_1 and those from V_2R_2. The violet fringes are produced by V_1 and V_2 while red fringes are produced by R_1 and R_2.

Suppose $V_1V_2 = d_1$ and $R_1R_2 = d_2$.

If $\frac{\lambda_V}{d_1} = \frac{\lambda_R}{d_2}$ the fringe width β will be the same and interference fringes due to different colours will overlap and white achromatic fringes are produced in the filed of view. The white and dark fringes are seen through the eyepiece or can be projected on a screen.

7.10 NON-LOCALIZED FRINGES

Point sources produce fringes, which can be seen at different distances from the source. As the screen is moved far from the source, the fringe spacing increases and conversely, when the screen is moved nearer to the source, the fringes come closer.

Therefore, we say the fringes are **non-localized**. Narrow sources produce non-localized fringes.

Figure : 24

7.11 VISIBILITIES OF FRINGES:

The contrast of the interference fringes can be quantitatively described by the parameter called visibility. Visibility is defined as

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots$$

(18)

the value of visibility varies between 0 and 1. When the fringes are of maximum intensity in the bright areas and totally dark in the dark areas, the visibility is equal to 1. As the phase difference increases, the coherence between the light waves decreases and the visibility is reduced. Finally, when the coherence between the two light waves disappears, I_{max} and I_{min} become equal and the visibility goes to zero. Then fringes are not observed and instead we observe uniform illumination.

From equ. (14.23) and (14.24), the visibility of the fringes produced by two beams can be expressed as

$$V = \frac{(E_1+E_2)^2-(E_1-E_2)^2}{(E_1+E_2)^2+(E_1-E_2)^2} = \frac{2E_1E_2}{E_1^2+E_2^2} = \frac{2\sqrt{I_1I_2}}{I_1+I_2} = \frac{2\sqrt{I_1/I_2}}{1+I_1/I_2} \quad \ldots\ldots\ldots\ldots$$

(19)

It is seen from the above equation that the closer the intensities of the two waves, the higher is visibility of the fringes. When $I_1 = I_2$, V = 1. It may be noted that V is always equal to 1 when monochromatic light is used.

7.12 FRINGE PATTERN WITH WHITE LIGHT

When the slit is illuminated with white light in any interference experiment, we obtain a central white fringe flanked by a few colored fringes. At the center of the screen O, there is zero path difference and the bright fringes produced by all the colors there, add to each other. As a result, a white fringe is produced at O. At other places away from O. the bright fringes in each order separate, because the fringe width is dependent on the wavelength ($\beta=\lambda D/d$). For example, the width of the red fringe is more than the blue fringe. Hence colored fringes are produced on either side of the central white fringe.

7.13 SUMMARY

When two light waves of same frequency and nearly some amplitude and having constant phase difference traverse in a medium and cross each other, there is redistribution in the intensity of light which is called interference of light.

Interference: Redistribution of energy due to superposition of waves.

Interference fringes: Pattern of dark and bright bands due to interference.

Superposition: Combining the displacements of two or more waves to produce a resultant displacement.

Coherence: Property of two or more waves with equal frequency and constant phase difference.

Coherent light: Light in which all wave trains have same frequency and its crests and troughs aligned in same directions which have constant phase difference.

Biprism: Combination of two prisms with their bases in contact.

Slit: A narrow opening for light.

7.14 TERMMINAL QUESTIONS

1. What is interference of Light?
2. Explain briefly Principle of Superposition.
3. State the fundamental conditions for the production of interference fringes.
4. Briefly discuss about fringe in Wave.
5. Write an essay on Interference of Light.
6. Write short notes on :
   (a) Fresnel's biprism
   (b) Lloyd's Single Mirror
   (c) Fresnel's Double Mirror
   (d) Achromatic Fringes
7. Discuss the experimental arrangement of Fresnel Biprism.
8. What is Young's double slit experiment? Find out the Position of bright fringes, dark fringes and fringe width.
9. Define visibility of Fringes.
10. Briefly explain fringe pattern with white light.

7.15 SOLUTION & ANSWER OF TERMINAL QUESTIONS

1. Section 7.3.1
2. Section 7.3
3. Section 7.3.5
4. Section 7.4
5. Section 7.3.1
6. (a) Section 7.6
   (b) Section 7.7
   (c) Section 7.8
   (d) Section 7.9
7. Section 7.6.1
8. Section 7.5
9. Section 7.11
10. Section 7.12

7.16 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.
2. Optics – Ajoy Ghatak
3. Optics – Principles and Application, K. K. Sharma, Academic Press.
4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.
5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT 08: INTERFERENCE BY DIVISION OF AMPLITUDES

STRUCTURE:

8.1 INTRODUCTION

In the previous unit, we discussed interference of light by division of wavefront wherein different portions of the wavefront propagate in different directions and then made to recombine. You may recall that in Young's double slit experiment, light coming out of a pin hole was allowed to fall on two holes, and the light waves emanating from these two holes interfered to produce the interference pattern. But interference of light responsible for the brilliant colour of thin oil films on water involves two light beams derived from a single incident beam by division of amplitude. When light is incident on a thin transparent plane-parallel film, the waves are reflected partially from the upper surface interfere with the waves partially reflected from the lower surface giving rise to beautiful colours. However, in this process, the amplitudes are reduced. This phenomenon is also responsible for colours in soap bubbles and cracks in glass. You may recall from the previous unit that in Lloyd's mirror, the interference takes place due to superposition of waves coming directly from the source and those reflected from an optically denser medium. In this arrangement, the interference pattern had a dark central fringe. This observation was explained by arguing that a phase change of took place when light waves were reflected at the surface of an optically "denser" medium. We begin this unit by giving the statement of the principle of reversibility of light. Later on we have discussed interference in thin transparent films with particular reference to fringes of equal inclination (Haidinger fringes) and fringes of equal thickness (Fizeau fringes). Further we discuss Interference by a wedge-shaped film. Lastly, we have discussed Newton's rings and how these rings are used to determine the wavelength of light.

8.2 OBJECTIVES

After studying this unit, you should be able to:

- Prove that when a light wave is reflected at the surface of an optically denser medium, it undergoes a phase change;
- Explain the origin of interference pattern produced by a parallel thin transparent film;
- Describe the formation, shape and location of interference fringes obtained from a thin wedge-shaped film;
- Discuss how Newton's rings are produced and used to determine wavelength of light;

❖ Explain why a thin coating of a suitable substance minimizes the reflection of light from a glass surface;

❖ Distinguish between fringes of equal inclination and fringes of equal thickness.

8.3 WHAT IS THE PRINCIPLE OF REVERSIBILITY OF LIGHT

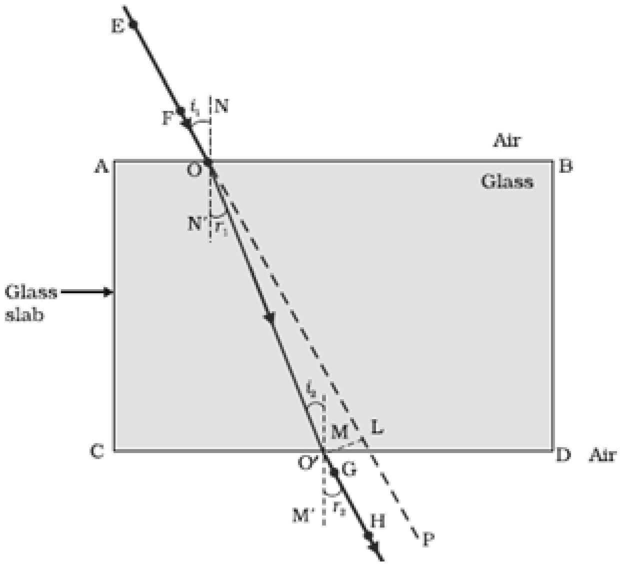


Figure : 1

Also, according to principle of reversibility of light, if the path of the light is reversed after suffering a number of reflections and refractions, then it retraces its path. This means that if a light ray travels from medium 1 to medium 2 and has angle of incidence and angle of refraction as i and r respectively, then if the light is incident from medium 2 at an angle r, then the angle of refraction in medium 1 will be i.

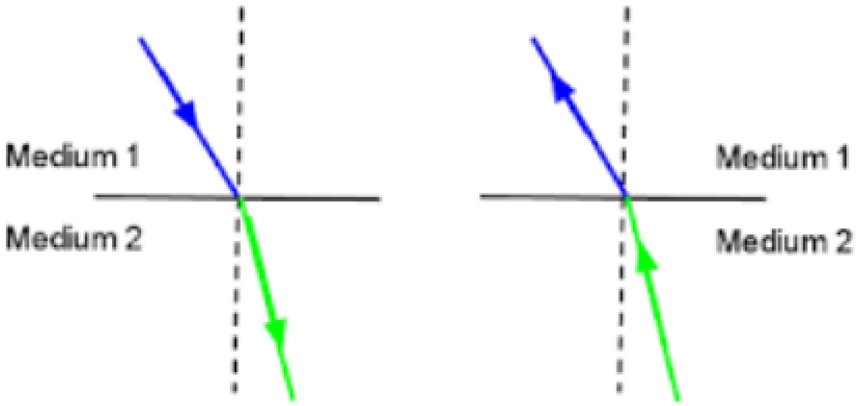
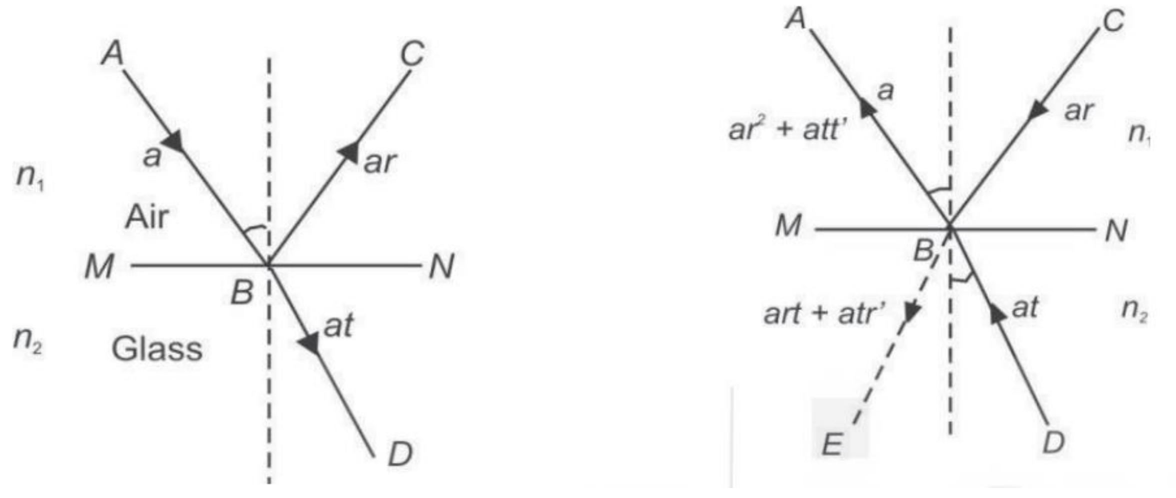


Figure : 2

8.2 STOKES' ANALYSIS OF PHASE CHANGE ON REFLECTION

From the preceding unit we have discussed that In Lloyd's mirror arrangement, the interference pattern had a dark central fringe. By arguing that a phase change of took place when light waves were reflected at the surface of an optically "denser" medium. Stokes used the principle of optical reversibility to verify this observation. The principle of optical reversibility states that a reflected or refracted light (ray) retraces its original path, if its direction is reversed, provided there is no absorption of light. Refer to Fig. 3(a), which shows the surface MN separating two optically different media (air and glass); glass being denser. We denote air as medium 1 and glass as medium 2.



(a) (b)

Fig. 3: **(a)** **Partial reflection and refraction of light at an air-glass interface;**

(b) **the optically reversed situation; the two rays in the lower left must cancel. In both cases $n_2 > n_1$, (n_1 and n_2 are the refractive indices of the media).**

An incident light wave, denoted by ray AB, is partly reflected along BC and partly transmitted (refracted) along BD. Suppose the amplitude of the incident wave AB is a, the fraction of the amplitude reflected is r, and the fraction transmitted when the wave is travelling from medium 1 to 2 is t. Then the amplitudes of the wave travelling along BC and BD will be ar and at, respectively.

Now, suppose that the directions of the reflected and transmitted (refracted) waves are reversed. As shown in Fig. 3(b), the wave BC, on reversal, gives a reflected wave along BA, and a transmitted wave along BE. You may now like to know the amplitudes of waves reflected along BA and transmitted wave along BE. You will agree that the amplitude of reflected wave along BA will be ar.r = ar^2 and the amplitude of transmitted wave along BE will be art. Further, the wave BD, on reversal, gives a transmitted wave along BA and a reflected beam along BE. Let us denote the fractions of amplitude reflected and transmitted when the wave is travelling from a denser medium (2) to a rarer medium (1) be r' and t'. This will contribute att' to the amplitude of the transmitted wave along BA and atr' to the amplitude of reflected wave along BE.

According to the principle of reversibility of light, the reflected and transmitted waves BC and BD, when reversed, should give rise to the original wave of amplitude a along BA. It means that there should be no wave along BE, i. e. component along BE should be zero. Similarly, the amplitude of the wave along BA should be equal to a. Mathematically, we write

$$art + atr' = 0$$
$$ar^2 + att' = a$$

equation 1 and 2

By rearranging terms in Eq. (1) & (2) we get

$$r' = -r$$
$$tt' = 1 - r^2$$

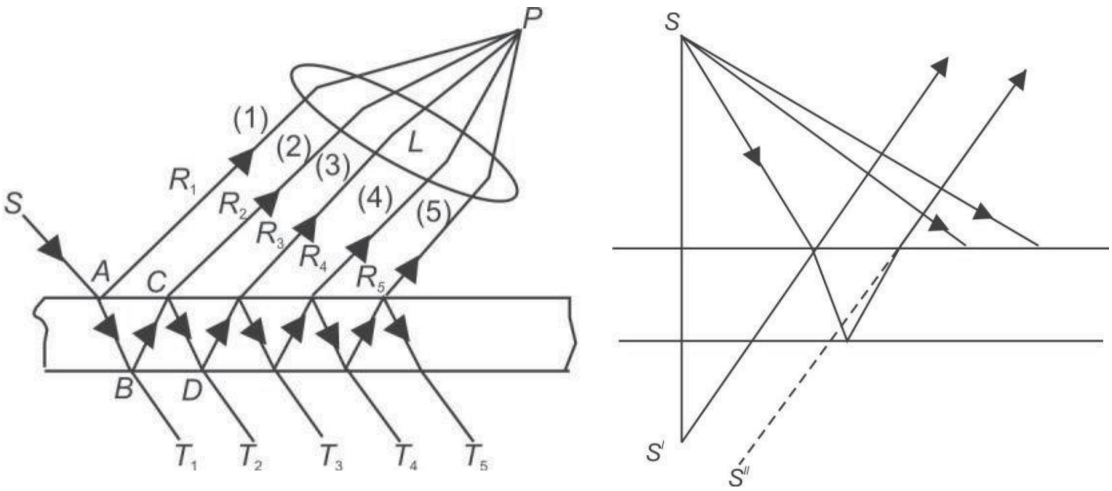
equation 3 and 4

These equations are known as Stokes' relations.

Eq. 3 shows that the fraction of amplitude reflected when incident wave is travelling from a rarer to denser medium (r) is numerically equal to the fraction of amplitude r' when incident wave is travelling from a denser to a rarer medium but have opposite sign. It means that these components are exactly out of phase with one another, i.e., their phase difference is π. Now two situations arise: No phase change occurs when a light wave is reflected by a denser medium but a phase change of occurs when a light wave is reflected by a rarer medium – and conversely, no phase change occurs when a light wave is reflected by a rarer medium then there must be a phase change of when a light wave is reflected by a denser medium. Now, out of these two alternatives, the second one is correct because it has been experimentally observed (180 or π on Lloyd's mirror) that the phase change of π occurs when light reaches the boundary from the rarer medium. Hence, light reflected by a material of higher refractive index than the medium in which the waves are travelling undergoes a phase change of (180 or π). Reflection by a material of refractive index lower than the medium in which the rays are travelling causes no phase change.

8.5 INTERFERENCE IN THIN FILMS

In Fig. 4(a). Light from a source S is incident on a thin film of soapy water, at A. A part of this will be reflected as ray (1) and part refracted along the direction AB. On



arrival at B, a part of the latter will be reflected to C, and a part refracted along BT_1. At C, the wave will again get partly reflected along CD and refracted as ray (2) along CR_2. Note that continuation of this process yields two sets of parallel rays, one on each side of the film. In each of these sets, of course, the amplitude of wave decreases rapidly in each step. Considering only the first two reflected rays (1) and (2) we observe that these two rays are in a position to interfere. This is because, if we assume S to be a monochromatic point source, the film serves as an amplitude-splitting device. Ray (1) and (2) may be considered as arising from the point source S due to reflection at the top and bottom surfaces of the thin film. But these two rays may respectively be considered as arising from two coherent virtual sources S' and S'' which are obtained by producing the two rays behind the film as given in fig 4 (b).

(a) **(b)**

Fig. 4: **(a) Multiple reflections in a soap film; (b) The interference pattern produced due to rays (1) and (2) is approximately the same as would have been produced by two coherent point sources S' and S''.**

If the set of waves (shown as parallel reflected rays) is now collected by a lens and focused at P, each wave would have travelled a different distance. So, the phase relationship between them may be such as to produce destructive or constructive interference at P giving rise to colours when seen by unaided eyes

8.5.1 PATH DIFFERENCE IN REFLECTED LIGHT

Path difference in reflected light is given in to Fig. 5. It In the given figure shows that light is incident on the thin film of soapy water at A at an angle i. Suppose the thickness of the film is t and refractive index of the material of the film μ is greater than one. At A, light is partly reflected along AR_1 giving the ray (1) and partly refracted along AB at an angle r. At B it is again partly reflected along BC and partly refracted along BT_1. Similarly, reflections and refractions occur at C. Since, the rays (1) and (2) have been derived from the same incident ray, they constitute a pair of coherent lights and interfere. Let CN and BM be perpendicular to AR_1 and AC. Since the paths of the rays AR_1 and CR_2 beyond CN are equal, the path difference between ray (1) and (2) is given by path ABC in film – path AN in air.

path difference = μ (AB + BC) – AN ……….. (5)

Note that,

$$AB = BC = \frac{BM}{\cos r} = \frac{t}{\cos r}$$

and $\quad AN = AC \sin i$

Since $AC = AM + MC$, we can write

$$AC = BM \tan r + BM \tan r$$

$$= 2t \tan r$$

Using this result in the relation for AN, we get

$$\therefore \quad AN = 2t \tan r \sin i$$

$$= 2t \frac{\sin r}{\cos r}(\sin i)$$

$$= 2t \frac{\sin r}{\cos r}(\mu \sin r) \qquad \left[\because \frac{\sin i}{\sin r} = \mu\right]$$

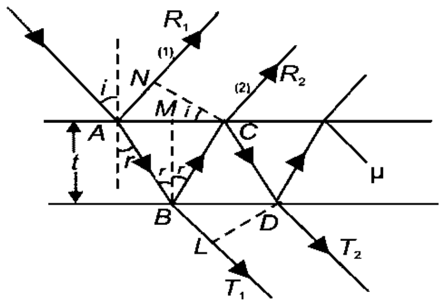


Fig. 5 Optical path difference between two consecutive rays arising in multiple reflections.

On substituting the values of AB, BC and AN in Eq. (5), we get

$$\text{path difference} = \mu\left(\frac{t}{cosr} + \frac{t}{cosr}\right) - 2\mu t \frac{sin^2 r}{cosr}$$

$$= \frac{2\mu t}{cosr}(1 - sin^2 r)$$

$$\boxed{Path\ difference = 2\mu t\ cosr} \qquad \ldots\ldots..$$

(6)

Recall that ray (1) undergoes a phase change of on reflection, while ray (2) does not, since it is internally reflected (6). So, we must include this fact in the expression for path difference obtained in Eq. (7.6). Since phase change of is equivalent to a path difference of $\frac{\lambda}{2}$, the effective path difference between rays (1) and (2) is given by path difference $= 2\mu t \, cosr + \frac{\lambda}{2}$ ………..

(7)

Note that the sign of phase change is immaterial.

We know that, if this path difference is an odd multiple of $\frac{\lambda}{2}$, we might expect rays (1) and (2) to be out of phase, and produce destructive interference. Thus, the condition for **destructive interference** for light reflected by a thin film can be expressed as

$$2\mu t \, cosr + \frac{\lambda}{2} = (2n+1)\frac{\lambda}{2}, \;\; where\; n = 0, 1, 2, \dots \dots$$

or $\qquad 2\mu t \, cosr = n\lambda$

……………….. (8)

Next, let us examine the phases of the rays, (3), (4), (5), … Since the geometry is the same, the path difference between rays (3) and (2) will also be given by Eq. (6). Since only internal reflections are involved in this case, the effective path difference will still be given by Eq. (6). Hence, if the condition given by Eq. (8) is fulfilled, ray (3) will be in the same phase as ray (2). The same holds true for all succeeding pairs, and so we conclude that, Eq. (8) expresses the condition that rays (1) and (2) will be out of phase, but rays (2), (3), (4), …, will be in phase with each other. Further, since ray (1) has considerably greater amplitude than ray (2), it is only logical to think that these will not annul each other completely. That is, the condition given by Eq. (8) may not produce complete darkness. But it is not so because addition of rays (3), (4), (5), …, which are all in phase with ray (2), gives a net amplitude which is just sufficient to make up for the difference and to produce complete darkness.

To prove that light reflected by a thin film can lead to complete darkness, refer to Fig. 6, which shows the amplitudes of successive rays when light undergoes multiple reflections.

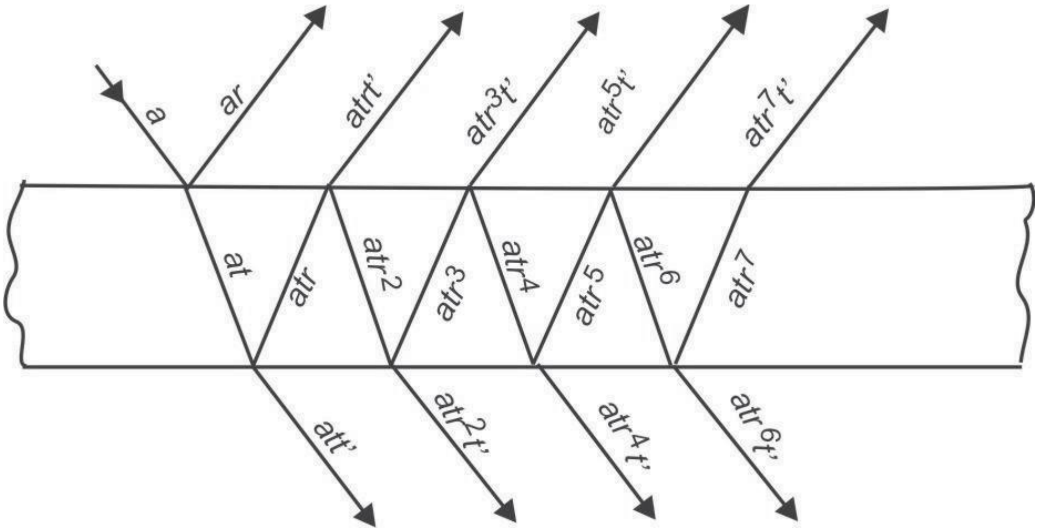


Fig. 6 Amplitude of successive rays when light undergoes multiple reflections.

On adding the amplitudes of all but the first reflected rays, on the upper side of the film, we obtain the resultant amplitude:

$$A = atrt' + atr^3t' + \ atr^5t' + atr^7t' + \cdots$$

$$= atrt'(1 + r^2 + r^4 + r^6 + \cdots)$$

Using series expression concept, for r less than one, it represents a geometrical series whose sum is finite, equal to1/(1-r^2), giving

$$A = atrt' \frac{1}{(1 - r^2)}$$

But from Eq. (4) we recall that = $1 - r^2$, . Using this result in the above relation, we get a compact relation for resultant amplitude of 2,3,4…. waves:

$$A = atrt' \left(\frac{1}{tt'}\right)$$

A = ar ……………… (9)

That is, the resultant amplitudes of all, but the first, reflected wave, on the upper side of the film, is just equal to the amplitude of the first reflected wave. Hence, we can conclude that when Eq. (8) is satisfied, there will be complete destructive interference. On the other hand, if the path difference given by Eq. (7) is an integral multiple of λ, waves denoted by rays (1) and (2) will be in phase and lead to **constructive interference**, i.e., when

$$2\mu t\ cosr - \frac{\lambda}{2} = n\lambda, \quad where\ n = 0, 1, 2, \ldots..$$

or $$2\mu t\ cosr = (2n+1)\frac{\lambda}{2} \qquad \ldots\ldots\ldots\ldots\ldots (10)$$

But rays (3), (5), (7), …, will be out of phase with rays (2), (4), (6), … Since (2) is more intense than (3), (4) is more intense than (5), etc., these pairs cannot cancel each other. As the stronger series combines with ray (1), the strongest of all, there will be maximum of intensity.

Thus, when a thin film is illuminated by monochromatic light, and seen in reflected light, it appears bright or dark according as $2\mu t\ cosr$ is an odd multiple of $\frac{\lambda}{2}$ or an integral multiple of λ, respectively.

$$\boxed{\begin{array}{ll} 2\mu t\ cosr = (2n+1)\frac{\lambda}{2} & (condition\ of\ maxima) \\ 2\mu t\ cosr = n\lambda & (condition\ of\ minima) \end{array}} \qquad \ldots\ldots. \quad (11(a) \quad \&$$

(b)

8.6 COLOURS IN THIN FILMS

When light is incident on a thin film, we receive rays of light reflected at its top and at the bottom surface. These rays interfere and depending on the path difference between the interfering rays, as given by Eq. (7), we note that it depends on t , the thickness of the film and r, the fraction of amplitude reflected r, and, hence, upon inclination of the incident rays. Note that inclination is determined by the position of the eye relative to the region of the film, which is being looked at. We know that sunlight consists of a continuous range of wavelengths (colours). At a particular point of the film, and for a particular position of the eye (i.e., for a particular t and a particular r), the rays of only certain wavelengths will have path difference, which satisfies the condition of maxima. Hence, only these wavelengths (colours) will be present with the maximum intensity. While some others, which satisfy the condition of the minima will be missing. Hence, the point of the film being viewed will appear coloured.

So far, we have considered viewing of thin film in reflected light. Suppose the eye is now situated on the lower side of the film, shown in Fig. 4 and Fig. 6. The light emerging from the lower side of the film can also be made to interfere.

What colours will arise?

SAQ : 1

A thin film of thickness 4×10^{-5} cmis illuminated by white light normal to its surface ($r = 0^o$). Its refractive index is 1.5. Of what colour will the thin film appear in reflected light?

SOLUTION The condition for constructive interference of light reflected from a film is

$$2\mu t\ cosr = (2n+1)\frac{\lambda}{2}, where\ n = 0, 1, 2, \ldots ..$$

Hence $\mu = 1.5; t = 4\times10^{-5}cm$ and $r = 0^o$ (since light falls normally) so that cosr = 1. On substituting the given values, we get

$$2\times1.5\times4\times10^{-5} = (2n+1)\ \frac{\lambda}{2}$$

or $\lambda = \frac{2\times2\times1.5\times4\times10^{-5}}{2n+1}$

$$\lambda = \frac{24\ \times10^{-5}cm}{2n+1} = \frac{24{,}000}{2n+1}\text{Å}$$

Taking n 0, 1, 2, ….. we get

$$\lambda = 24000\ \text{Å}, 8000\ \text{Å}, 4800\ \text{Å}, 3431\ \text{Å}, \ldots.$$

These wavelengths are reflected most strongly. Of these, the wavelength lying in the visible region is 4800 Å which corresponds to blue colour.

Let us now discuss, as to which colours will arise when the film is viewed in refracted (transmitted) light. As before, we first calculate the path difference between the transmitted waves, shown as rays in Fig. 4.

The path difference between the transmitted rays BT_1 and DT_2 is given by

$$\mu(BC + CD) - BL = 2\mu t\ cosr$$

Note that in this case, there is no phase change due to reflections at B and C because in both cases light is reflected at the denser-rarer medium interface from the denser side. Hence, the effective path difference between BT_1 and DT_2 is also $2\mu t\ cosr$.

The two rays BT_1 and DT_2 will reinforce each other, if

$2\mu t\ cosr = n\lambda \qquad (conditin\ of\ maxima)$ ……… (12a)

where n = 1, 2, 3 ……..

In this case, the film will appear brighter in the transmitted light.

The two rays will annihilate the effect mutually if

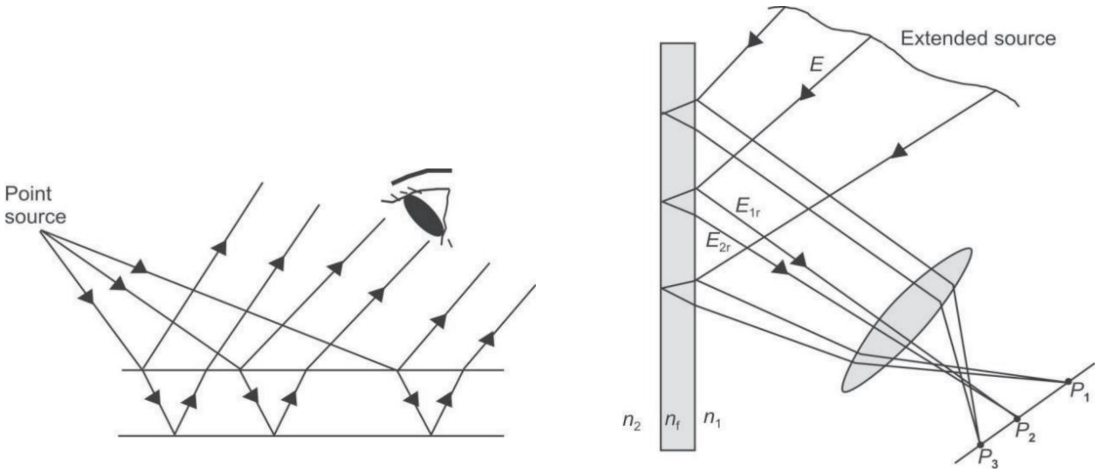
$2\mu t\ cosr = (2n + 1)\ \frac{\lambda}{2} \qquad (conditin\ of\ minima)$ …….. (12b)

where n = 0, 1, 2, …….. and the film will appear dark in transmitted light.

Now, reflect on the results that we have obtained in this section. Eqs. (11a), (11b), (12a) and (12b), It shows that the conditions for the maxima and minima for the reflected light are just reverse of those in transmitted light. It means that only those colours will be visible in transmitted light which are missed in reflected light. In other words, the appearances of colours in the two cases are complimentary. Hence, if the film appears bright in reflected light, it will appear dark in transmitted light and vice versa.

8.7 HAIDINGER FRINGES (FRINGES OF EQUAL INCLINATION)

If a small aperture lens is used to focus light reflected by a thin film and incident from



a point source, the interference fringes will be visible only from a small portion of the film. This is because light from only a small part of the film will reach the lens and be seen (Fig. 7a). However, if an extended source is used, light will reach the lens from various directions, and the fringe pattern will spread out over a larger area of the film, as shown in Fig. 7b.

(a) **(b)**

Fig. 7: **Fringes are respectively seen from a small portion of the film and a larger region of the film if the source is a) a point source, and b) an extended source.**

Remember, the angle of incidence i or equivalently angle of refraction r, controls the path difference. The fringes appearing at point P_1 and P_2 in Fig. 8 are, accordingly, known as **fringes of equal inclination.**

It may be noted that as the film becomes thicker, the separation AC in Fig. 5 between rays (1) and (2) also increases, since AC = 2t tan r. When only one of the two rays enters the eye, the interference pattern disappears. However, the larger lens of a telescope should then be used to gather both rays so that the pattern becomes visible. The separation between rays (1) and (2) can also be reduced by reducing i, i.e., by viewing the film at nearly normal incidence.

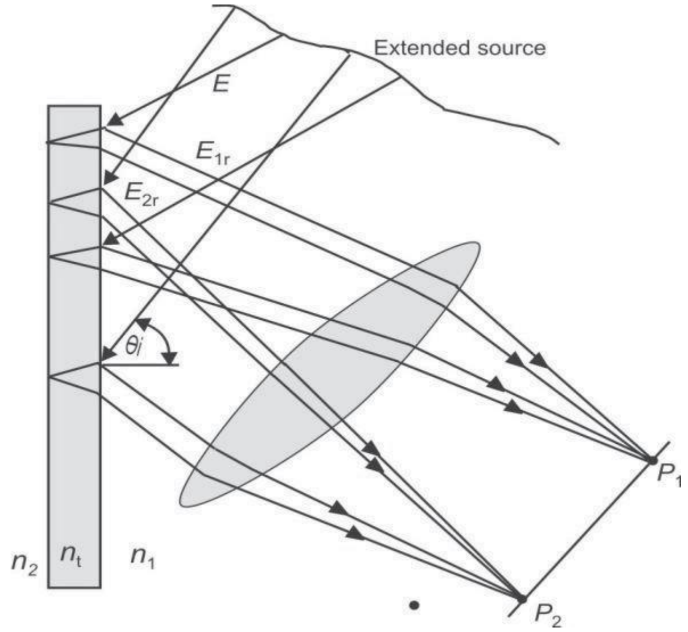


Fig. 8: **Haidinger fringes: All rays inclined at the same angle arrive at the same point.**

The equal inclination fringes obtained in thick plates are known as Haidinger fringes. With an extended source, the symmetry of the set up requires that the interference pattern consists of a series of concentric circular bands centred on the perpendicular drawn from the eye to the film, as shown in Fig. 9.

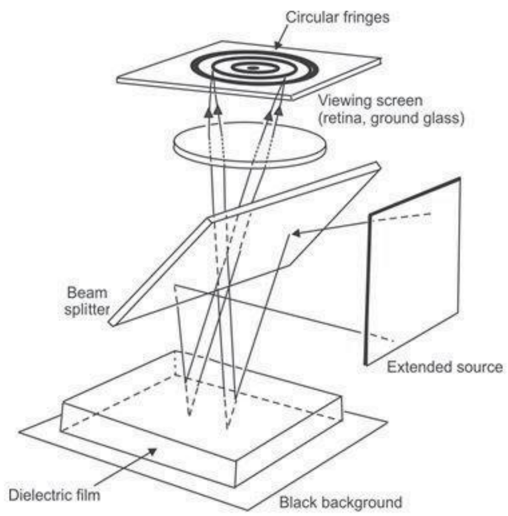
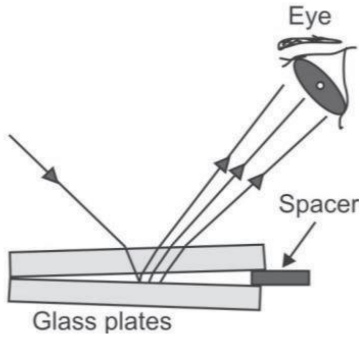


Fig. 9: **Circular Haidinger fringes centred on the axis of the lens.**

Haidinger fringes are formed at infinity and are observed using a telescope focussed at infinity. Such fringes are observed in Michelson interferometer.



8.8 FIZEAU FRINGES (FRINGES OF EQUAL THICKNESS)

Interference fringes, for which thickness of the film rather than the angle of refraction is the dominating parameter, are referred to as **fringes of equal thickness**. The film is not plane- parallel i.e. the film has continuously varying thickness. When such a film is illuminated by a monochromatic light source than equidistant straight line interference fringes parallel to the thin edge of the wedge are formed. Each such fringe is the locus of all points in the film for which thickness is a constant. Such fringes are localised on the film itself and are observed by a icroscope focussed on the film. Fringes due to the wedge-shaped film belong to this class.

Fringes of equal thickness can be distinguished from Haidinger fringes by the manner in which the diameters of the rings vary with order n. The central region in the Haidinger pattern corresponds to the maximum value of n, whereas just the opposite applies to fringes of equal inclination.

8.9 INTERFERENCE BY A WEDGE SHAPED FILM

So Far, we have assumed the film to be of uniform thickness. However, many a time, the fringes are produced by films of varying thickness, i.e., the film is not plane-parallel. Such films find practical application for testing optical surfaces for flatness.

We now discuss the interference pattern produced by a wedge-shaped film, whose plane surfaces are not parallel (Fig. 10(a) and 10(b)). The interfering rays do not enter the eye parallel to each other but seem to diverge from a point near the film.

(a) **(b)**

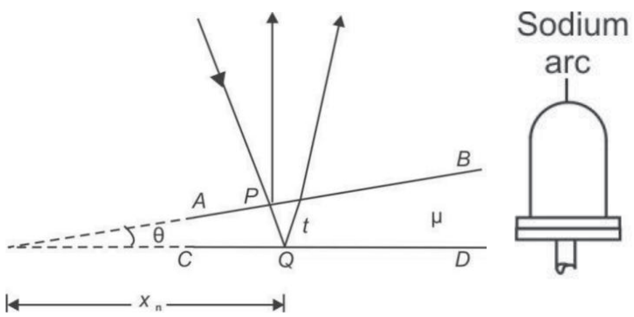


Fig. 10: **Fringes of equal thickness: a) method of visual observations; b) a parallel beam of light incident on a wedge-shaped film.**

To obtain conditions for maxima and minima of the interference pattern produce by a thin wedge-shaped film of refractive index μ.

Let us consider a film bounded by two plane surfaces AB and CD, inclined at an angle as shown in Fig. 10(a). When produced backwards, the lines corresponding to these surfaces meet at point O. The thickness of the film increases from A to B. Suppose the film is illuminated by a monochromatic source of light taken in the form of a slit held parallel to the edge of the wedge and light is incident perpendicularly on the surface AB. (The edge is the line passing through O and normal to the plane of the paper.) Interference occurs between the rays emerging from the upper and lower surfaces of the film. In this case, the path difference for a given pair of rays is essentially the same as that given by Eq. (6). But, if it is assumed that wedge angle is small, and since the light is assumed to be incident almost normally at a point P on the film, the factor cos r may be taken to be equal to unity. Then, the path difference between the rays reflected at the upper and the lower surfaces will be same, equal to $2\mu t$, where t is the thickness of the film at P.

Recall that an additional path difference of $\frac{\lambda}{2}$ is introduced for the light reflected from the upper surface. So, the effective path difference between the two reflected rays will be

$$2\mu t - \frac{\lambda}{2} \qquad \text{………… (13)}$$

and the condition for bright fringes can be expressed as

$2\mu t - \frac{\lambda}{2} = n\lambda$

or $\quad 2\mu t = (2n-1)\frac{\lambda}{2}$ ……….. (14)

Similarly, the condition for dark fringe is given by

$2\mu t = n\lambda$ ………… (15)

From these results, it is clear that to observe a bright or dark fringe of a particular order for a wedge shaped film with small wedge angle and normal incidence of light, t must be constant. Note that in a wedge-shaped film, t remains constant only along lines parallel to the thin edge of the wedge. Therefore, the bright and dark fringes are obtained as straight lines parallel to the thin edge of the wedge. At the thin edge, t = 0 and therefore path difference = $\frac{\lambda}{2}$ which corresponds to the condition of minimum intensity. It means that the edge of the wedge-shaped film will be dark. The resulting fringes resemble the localized fringes in the Michelson interferometer and appear to be formed in the film itself.

8.9.1 Fringe Width

We know that, a wedge shaped film with small wedge angle gives rise to an interference pattern comprising dark and bright straight line fringes when light is incident almost normally. The nth dark fringe, we have

$$2\mu t = n\lambda$$

Suppose this fringe is obtained at a distance x_n from the thin edge. Then $t = x_n \tan\theta = x_n\theta$ (when θ is small). Hence, for the nth dark fringe, we can write

$2\mu\, x_n\theta = n\lambda$ …….. (16)

Similarly, if the (n + 1)th dark fringe is obtained at a distance x_{n+1} from the thin edge, then we can write

$2\mu\, x_{n+1}\theta = (n+1)\lambda$ ………….. (17)

On subtracting Eq. (16) from Eq. (17), we get

$2\mu\theta\ (x_{n+1} - x_n) = \lambda$

Hence, we obtain the following expression for fringe width (β) :

$$\beta = x_{n+1} - x_n = \frac{\lambda}{2\mu\theta} \quad \text{................ (18)}$$

It is measured in radians. Such fringes are referred to as **fringes of equal thickness**. This result shows that fringe width depends on the wavelength of incident light, the refractive index of the film and the angle of wedge. It means that if we use mercury lamp, instead of a sodium lamp, to observe interference pattern, it will have coloured fringes.

8.10 NEWTON'S RINGS

A Ring of Newton's is an Optical phenomenon that appears as a Ring of dark-colored lights or bands when one piece of glass is Convex and rests on another piece of flat glass. Therefore, there is space between them that is filled with air. Light waves are said to cause this phenomenon as their interference results in a brightening of light when their crests coincide, but the opposite occurs when the crest meets the trough. This is described as the destruction of light. During the transmission of light between the two pieces of glass, waves from both the top and bottom surfaces of the air film interfere with each other. This section will present all the information related to Newton's Ring. Newton was the first Physicist to investigate the Rings quantitatively, and he is credited with naming them after him.

It is said that Newton's Rings are formed by the interference pattern between two surfaces caused by the light reflecting between them. There are two flat surfaces on either side of the sphere. In 1704, Isaac Newton described the effect in a treatise called Opticks based on his research. Newton's Rings are visible when viewed with monochromatic light as alternating bright and dark circles located at the point of contact between the two surfaces. The different wavelengths of light conflict at very different levels of thickness in the layer of air between the two surfaces when viewed in white light. This results in a pattern of a concentric circle of Rainbow colors.

An Optical glass that has a flat surface is placed on a very slightly curved Convex glass to create the pattern. At other points, there is a slight air gap between the two surfaces of the two pieces of glass which are in contact only at the center. The increasing distance between the center and the microscope is referred to as the radial distance. On the right side of one of the pieces, there is a small gap growing from left to right. In the monochromatic example, the source of color is a monochromatic single source of light shining through the top piece and refracting off the bottom and top surfaces. In the resulting superposition, the two rays are combined.

This ray, however, generally travels a relatively long path as it reflects off the surface of the bottom. Two times the gap between the surfaces equals the additional path length.

A 180° phase reverse happens when the ray reflects off the bottom piece of glass. As a result, the phase reversal due to the internal reflection of the other ray from the underside of the top glass does not occur.

8.10.1 Principle of Newton's Ring

Newton's rings is a phenomenon in which an interference pattern is created by the reflection of light between two surfaces; a spherical surface and an adjacent touching flat surface. It is named after Isaac Newton, who investigated the effect in 1666.

8.10.2 Theory of Newton's Ring

When a plano-convex lens with its convex surface is placed on a plane glass plate, an air film of gradually increasing thickness is formed between the lens and the glass plate. The thickness of the air film is almost zero at the point of contact O and gradually increases as one proceeds towards the periphery of the lens. If monochromatic light is allowed to fall normally on the lens, and the film is viewed in reflected light, alternate bright and dark concentric rings are seen around the point of contact. These rings were first discovered by Sir Isaac Newton, hence named as **Newton's Rings**. If it is viewed with the white light then coloured fringes are obtained. The experimental arrangement of the Newton's Ring apparatus is shown in *figure 11*.

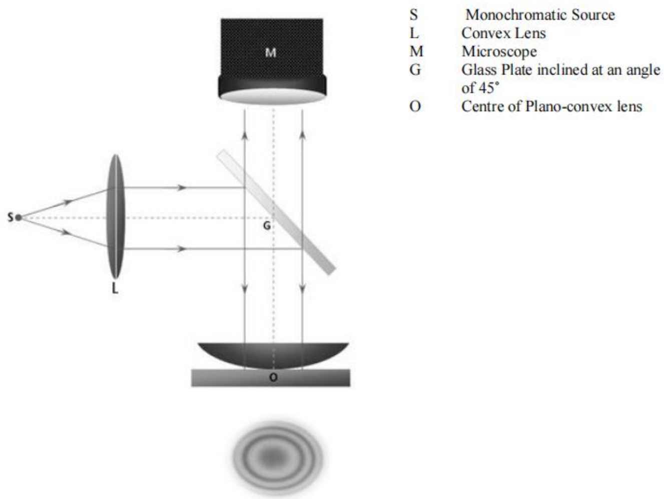


Fig. 11: Experimental arrangement of the Newton's Ring apparatus

A parallel beam of monochromatic light is reflected towards the lens *L*. Consider a beam of monochromatic light strikes normally on the upper surface of the air film. The beam gets partly reflected and partly refracted. The refracted beam in the air film is also reflected partly at the lower surface of the film. The two reflected rays, i.e. produced at the upper and lower surface of the film, are coherent and interfere constructively or destructively. When the light reflected upwards is observed through microscope *M* which is focused on the glass plate, a pattern of dark and bright concentric rings are observed from the point of contact *O*. These concentric rings are known as *Newton's Rings*.

The path difference between the two successive reflected rays are obtained using wedge shaped interference case as:

$$\Delta = 2\mu t \cos(r+\theta) \pm \lambda/2$$ ……………. (19)

For air film, $\mu = 1$ and at normal incidence, $i = r = 0$. Since, θ is very small then $\cos\theta = 1$.

Hence, eq. (19) reduces to

$$\Delta = 2t \pm \lambda/2$$

At the point of contact of the lens and the glass plate (O), the thickness of the film is effectively zero i.e. $t = 0$

$$\Delta = \pm\lambda/2$$

This is the condition for minimum intensity. Hence, the centre of Newton rings generally appears dark.

8.10.3 Diameter of Newton's Ring

Diameter of Bright Rings :

Consider a ring of radius *r* due to thickness *t* of air film as shown in the *figure 12.*

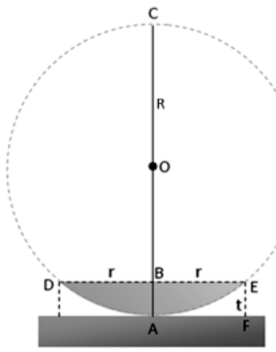


Fig. 12.

According to the geometrical theorem (i.e. property of the circle), the product of intercepts of the intersecting chord is equal to the product of sections of the diameter.

$$\overline{DB} \times \overline{BE} = \overline{AB} \times \overline{BC}$$

$$r \times r = t(2R - t)$$

$$r^2 = 2Rt - t^2$$

Since t is very small hence t^2 will also be negligible, thus,

$$r^2 = 2Rt$$

$$t = \frac{r^2}{2R} \qquad \text{………………} (20)$$

a. Condition for a bright ring (constructive interference in thin film)

$$2\mu t = (2n-1)\frac{\lambda}{2} \qquad \text{…………………..} (21)$$

Where n = 1, 2, 3 …….

Putting eq. (20) in eq. (21) we get

$$2\mu\left(\frac{r^2}{2R}\right) = (2n-1)\frac{\lambda}{2}$$

Radius of the n^{th} bright ring becomes

$$r_n^2 = (2n-1)\frac{\lambda R}{2\mu}$$

Thus diameter of the n^{th} bright ring is

$$\left(\frac{D_n}{2}\right)^2 = (2n-1)\frac{\lambda R}{2\mu}$$

$$D_n^2 = 2(2n-1)\frac{\lambda R}{\mu}$$

$$D_n = \sqrt{2(2n-1)\frac{\lambda R}{\mu}} \qquad \text{……………..} (22)$$

If the medium considered is air then μ=1and eq. (22) simplifies to

$$D_n = \sqrt{2(2n-1)\lambda R}$$

$$D_n \propto \sqrt{(2n-1)} \qquad \text{where } n = 1, 2, 3\ldots..$$

Thus, ***diameter of the bright rings is proportional to the square root of odd natural numbers***.

b. Condition for a dark ring (destructive interference in thin film)

$$2\mu t = n\lambda \qquad \text{where } n = 0, 1, 2, 3\ldots..$$

Putting eq. (20) in eq. (23) we get

$$2\mu \frac{r^2}{2R} = n\lambda$$

Radius of the n^{th} dark ring becomes

$$r_n^2 = \frac{n\lambda R}{\mu}$$

Thus, diameter of the n^{th} dark ring is

$$\left(\frac{D_n}{2}\right)^2 = \frac{n\lambda R}{\mu}$$

$$D_n^2 = \frac{4n\lambda R}{\mu}$$

$$D_n = \sqrt{\frac{4n\lambda R}{\mu}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots. (24)$$

where n = 0, 1, 2, 3 ……

If the medium considered is air then μ = 1 and eq. (24) simplifies to

$$D_n = \sqrt{4n\lambda R}$$

$$D_n \propto \sqrt{4n\lambda R}$$ where $n = 0, 1, 2, 3\ldots.$

Thus ***diameter of the dark rings is proportional to the square root of the natural numbers***.

8.10.4 Applications of Newton's Ring

1. Determination of Wavelength of Monochromatic Light (λ)

The diameter of the nth dark ring is given by:

$$D_n^2 = 4n\lambda R$$ ………………

(25)

Similarly, the diameter of the (n + p)th dark ring is given by:

$$D_n^2 = 4(n+p)\lambda R$$ ……………….

(26)

Subtracting eq. (25) from eq. (26), we get

$$\lambda = \frac{D_{n+p}^2 - D_n^2}{4pR}$$

where, is an integer.

2. Determination of Refractive Index of the Liquid (μ)

The diameter of the nth dark ring in air film is given by :

$$\left(D_n^2\right)_{air} = 4n\lambda R$$

Similarly, the diameter of nth dark ring in liquid film is given by :

$$\left(D_n^2\right)_{liquid} = \frac{4n\lambda R}{\mu}$$

Therefore, the Refractive Index of the Liquid is obtained as:

$$\mu = \frac{\left(D_n^2\right)_{air}}{\left(D_n^2\right)_{liquid}}$$

8.11 MICHELSON INTERFEROMETER

The Michelson interferometer is the best example of what is called an amplitude-splitting interferometer. It was invented in1893 by Albert Michelson, to measure a standard meter in units of the wavelength of the red line of the cadmium spectrum. With an optical interferometer, one can measure distances directly in terms of wavelength of light used, by counting the interference fringes that move when one or the other of two mirrors are moved. In the Michelson interferometer, coherent beams are obtained by splitting a beam of light that originates from a single source with a partially reflecting mirror called a beam splitter. The resulting reflected and transmitted waves are then re-directed by ordinary mirrors to a screen where they superimpose to create fringes. This is known as interference by division of amplitude. This interferometer, used in 1817 in the famous Michelson- Morley experiment, demonstrated the non-existence of an electromagnetic-wave-carrying ether, thus paving the way for the Special theory of Relativity.

An interferometer is an instrument in which the phenomenon of interference is used to make precise measurements of wavelengths or distances.

8.11.1 Principle of Michelson Interferometer

In Michelson interferometer, a beam of light from an extended source is divided into two parts of equal intensities by partial reflection and refraction. These beams travel in two mutually perpendicular directions and come together after reflection from plane mirrors. The beams overlap on each other and produce interference fringes.

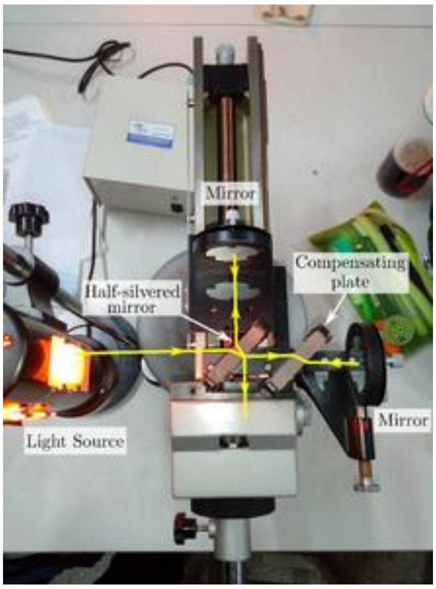


Figure 12

8.11.2 Construction of Michelson Interferometer

It consists of a semi silver plate P placed at an angle of 45 degrees to the horizontal. M1 and M2 are two highly polished silver mirrors. These Mirrors are provided with a labeling screw at the back.

Mirror M1 can be moved towards or away from plate P with the help of micrometer screws. P1 is another glass plate whose thickness is exactly equal to the thickness of plate P.

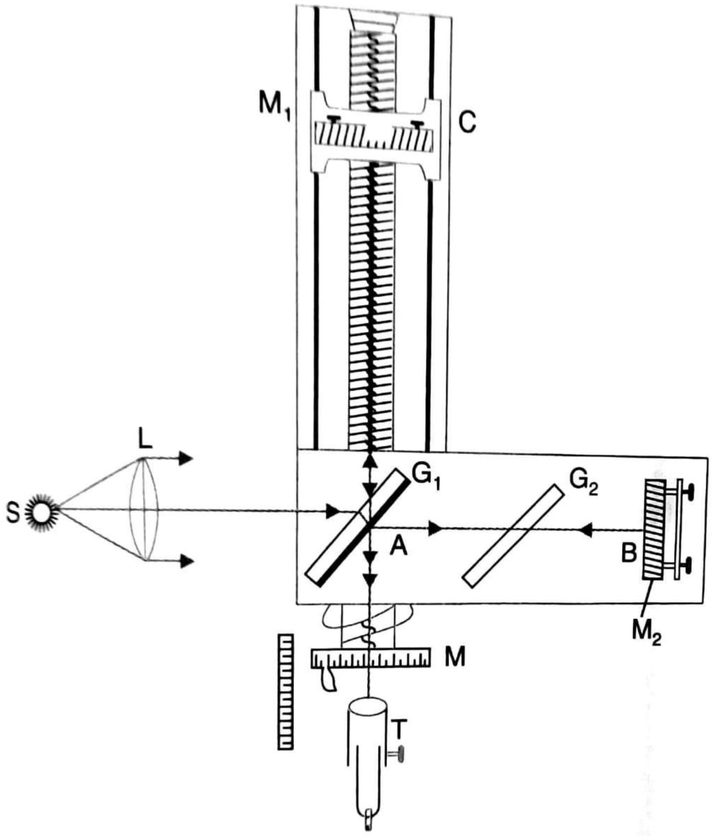


Figure 13

It is transparent and parallel to P. S is a monochromatic source of light, say sodium lamp. Lens L is a Converging lens whose function is made to the source of light as an extended one.

8.11.3 Working of Michelson Interferometer

Light from an extended monochromatic source S rendered parallel by the lens L is made to fall on semi-silvered plate P. Here, the light is divided into a reflected and transmitted beam of equal intensity.

The reflected beam which falls on mirror M1 and the transmitted beam that falls on mirror M2 is incident normally. When the Apparatus is in normal adjustment.

After reflection from mirror M1 and M2, a part of the amplitude of the wave from M1 is transmitted along with AE. While a part of the amplitude of the wave from M2 is reflected along with AE.

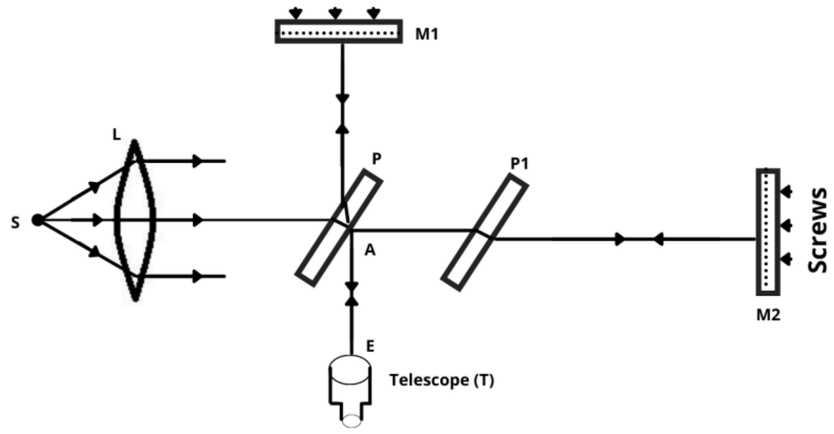


Figure : 14

As these two waves entering the telescope are derived from the same source S, hence these waves are coherent waves.

Coherent waves interfere with each other and interference fringes formed and seen through the telescope. The ray which gets reflected from M1 crosses the plate P twice while the ray which from M2 Travels only in the air.

This means that the ray which travels along the mirror M1 covers an additional part 2 (n-1) t, where t is the thickness of the plate P and n is its refractive index.

Which does not get any difficulty if we are using monochromatic light but if white light is used. Then it does create a problem because of the variation of n with wavelength.

To overcome this difficulty another plate P1 is introduced between P and mirror M2. This plate P1 is called a compensating plate.

The plate P and P1 are of equal thickness being cut from a single optical plane parallel plate to ensure the equality of thickness and the nature of the material.

The function of the compensating plate is that the Ray of light traveling towards M1 and M2 must travel equally **pass through the glass plate.**

The phase changes on reflection at mirror M1 and M2 are smaller, the phase changes due to reflection in air and class are also similar, is equal to $\boldsymbol{\pi}$**.** The two rays reaching the telescope interfere constructively or destructively depending upon the path difference.

The path difference between the two rays reaching the telescope is m**λ. Where M is the integer and λ is the wavelength of light used.** Then constructive interference takes place.

On the other hand, if the path difference is (2m+1) **λ/2,** Then destructive interference occurs. The path difference between the two beams AM1 and AM2 can be altered by moving mirror M1.

8.11.4 Conditions for different Shape of Fringe

When we see through the telescope, the mirror M1 is seen directly and a virtual image of M2 is also seen in M1. If the distance of M1 and M2 from plate P are not exactly equal, the image of M2 is formed slightly ahead or at the back of M1.

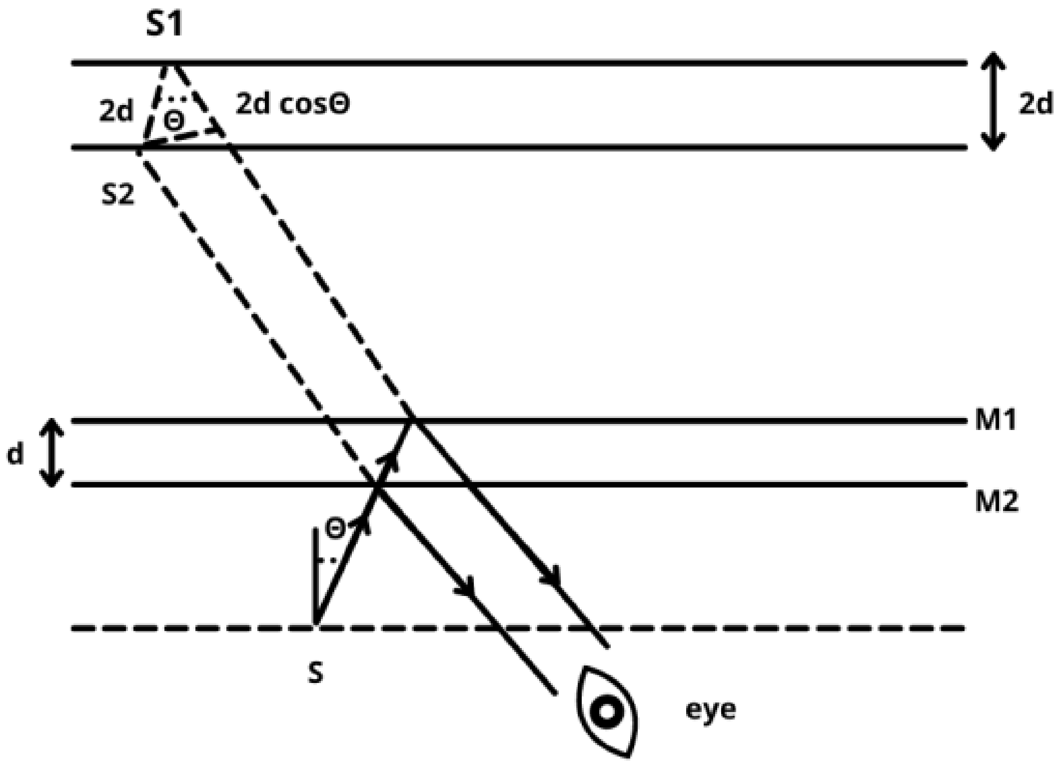


Figure : 15

Thus, the system reduces to the equivalent of an air film and enclosed between the mirror M1 and the virtual image of M2.

Now to see whether both the mirrors are at right angles to each other or not, a pin is placed between the source as and the plate P. Looking along EA, two images of the pin are seen.

The Adjustment of mirror M2 is made with the help of three screws such that both the images of the pin coincide. Which means both the mirrors are at right angles to each other.

The pin is now removed. Now if the distance AM1 and AM2 are exactly equal the field of view shall be totally dark. A slight movement of mirror M1 forward or backward shall produce concentric circular fringes in the field of view.

8.11.5 Types of fringes in Michelson Interferometer.

The shape of the fringes formed in the michelson interferometer depends on the inclination between mirror M1 and M2. Let M2 be the virtual image of mirror M1 as S' be the image of source S.

Let D be the separation between mirror M1 and the virtual image M2' of the mirror M2. Therefore, the interference pattern formed will be due to air film enclosed between M1 and M2'. Let S1' and S2' be the images of the Saras M1 and M2 respectively.

It means, we have two coherent sources S1' and S2' to obtain due to the process of the division of amplitude.

8.11.6 Circular Fringes

If the mirror M1 and M2 are exactly at right angles, then the mirror ember and the image M2' two of the mirror M2 are parallel to each other.

Then circular fringes are formed. Let D be the distance of mirror M1 and mirror M2 from the plate P then the distance between M1 and M2' who is also d. It means, the thickness of air film enclosed between M1 and M2' is d as shown in the figure.

However, the distance between two coherent sources S1' and S2' is 2d. The path difference between the rays coming from S1' and S2'=2d cos**θ.** Where **θ** is the angle of inclination of the ray Falling On The Mirror.

8.11.7 Localised Fringes

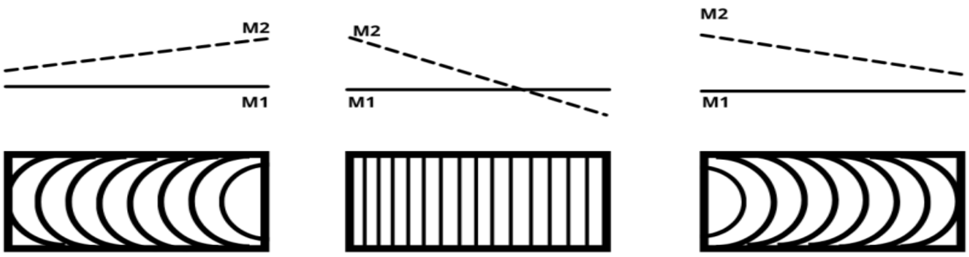


Figure : 16

When mirrors M1 and M2' are not exactly parallel or not equally distant, a wedge-shaped film is formed between them. The path of two reflected rays, originating from the same incident ray why reflection from M1 and M2' you are no longer parallel.

They intersect near M1 and hence the fringes are formed near M1. The fringes are called localized fringes and to see them the eye must be focused on the vicinity of M1.

These fringes are curved with their convex side toward the thin edge of the wedge as shown in the figure. The thin edge of the wedge is to the left and therefore the observer fringes are convex toward the left.

As we go on decreasing the separation between M1 and M2', the fringes move across the field of view away from the thin as of the wedge and at the same time gradually become straight. When M1 and M2' intersect, the lines are perfectly straight as shown in the figure.

We have to wedges opposing each other so the line should appear curved on both sides of the intersection but for a small field of view, they appear straight. When M1 is still moved such that the mirror M1 and the virtual image M2’ of mirror M2 get a position as shown in the figure.

The fringes are again curved but with their convex side towards the right. Localized fringes become invisible for large path differences of the order of several millimeters.

8.11.8 White Light Fringes

If monochromatic light is replaced by white light, its constituent wavelength gives rise to its own set of fringes two different widths. The zero-order fringes corresponding to each wavelength will coincide and hence we get a dark central fringe.

But if the path difference between the interfering rays is considerable. The central fringe will be surrounded by a few colored fringes. And then there will be so much overlapping that the pattern appears to be white.

The importance of these white light fringes is that the position of zero-order fringe which is dark can be located very easily in the Michelson interferometer.

8.11.9 Applications of Michelson Interferometer

Michelson Interferometer is used to determine

- Wavelength of monochromatic light.
- The refractive index of a thin film.
- Resolution of spectral lines.
- The evolution of meters in terms of the wavelength of light.
- The angular diameter of stars.
- Presence of ether.
- The accuracy of the surface of the prism and lens.

8.11.10 Comparison of Newton's Ring and Michelson Interferometer Fringes

Newton's Ring	Michelson's Interferometer
$\boxed{2\mu t \cos r = n\lambda}$ $here, t$ –variable and r-fix	$\boxed{2d \cos\theta = n\lambda}$ d-fix and θ-variable (angle of incidence)
Circular fringes because of locus of equal thickness of film.	Circular fringes because of equal inclination of light of incidence.
Newton's ring have minimum order at the centre.	Michelson's interferometer have maximum order at the centre.
It is viewed with the help travelling Microscope because they are localized in the plane of wedge shaped air film.	It is viewed with the help of Telescope.
It is also known as Hadinger fringes (Equd Indiviation)	It is also know as Fizeau Fringes.

8.11.11 Similarities of Newton's ring and Michelson Interferometer Fringes.

The Newton's ring is a phenomenon wherein light gets reflected between two surfaces to create an interference pattern. In Michelson interferometer, light is reflected back on beam splitter to create an interference pattern.

Similarities :

1. Both phenomenons are used to create interference patterns.
2. In both the phenomenons, light beams are reflected upon surfaces to create interference patterns.

8.12 WHAT IS FEBRY-PEROT INTERFEROMETER

The Fabry-Perot interferometer uses the phenomenon of multiple beam interference that arises when light shines through a cavity bounded by two reflective parallel surfaces. Each time the light encounters one of the surfaces, a portion of it is transmitted out, and the remaining part is reflected back. The net effect is to break a single beam into multiple beams which interfere with each other. If the additional optical path length of the reflected beam (due to multiple reflections) is an integral multiple of the light's wavelength, then the reflected beams will interfere constructively. More is the number of reflections inside the cavity, sharper is the interference maximum. Using Fabry-Perot (FP) interferometer as a spectroscopic tool, concepts of finesse and free spectral range can be understood.

8.12.1 Principle of Febry-Perot Interferometer

It is based on principle of interference by Multiple reflections.

In Fabry-Perot inerferometer both reflected and transmitted beams interfere with each other, so it is usually used in the transmissive mode. In case of Fabry-Perot Interferometer, each wavelength produces its own ring pattern and the patterns are spearated from each other. Therefore, Febry-Perot Interferometer is suitable to study the fine structure of spectral lines.

8.12.2 Working of Febry-Perot Interferometer

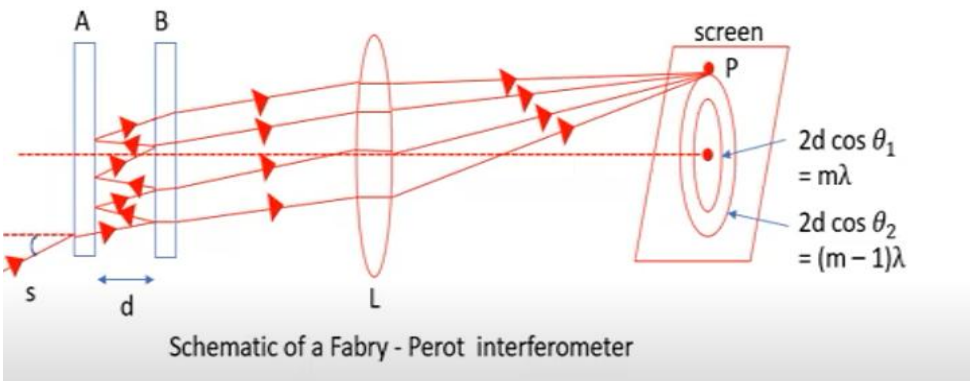


Figure : 17

- If S is a source of monochromatic light.
- Let L_1 is a convex lens (not shown in fig) which makes the rays parallel.
- An incident ray surffers a large number of internal reflection successively at the two silvered surfaces.
- At each reflection a small fraction of the light is transmitted also.

- Thus, each incident ray produces a group of coherent and parallel transmitted rays with a constant path difference between any two successive rays.
- A second convex lens L brings these rays together to a point in its focal plane where they interfere.
- Hence, the rays from all points of the source produce an interference pattern on a screen placed in the focal plane of L.

8.12.3 Construction of Febry-Perot Interferometer

The Fabry-Perot interferometer essentially consists of two glass plates A and B separated by a distance (Fig.). The inner surfaces of the plates are coated with a thin film of aluminum which reflected about 75% of the incident light. The plate B facing the observer is fixed and is provided with screws with which the reflecting surface of B can be made parallel to that of A. the plate A is mounted on a carriage which can be moved in a direction perpendicular to the reflecting surfaces by means of an accurate screw so that the thickness of air film between the coated surfaces of the plates A and B can be varied. Light from monochromatic extended source rendered parallel by collimating lens L, suffers multiple reflections in the air film between plates A and B. The transmitted light interfere and circular fringes of equal inclination are formed at the focal plane of the objective of the telescope.

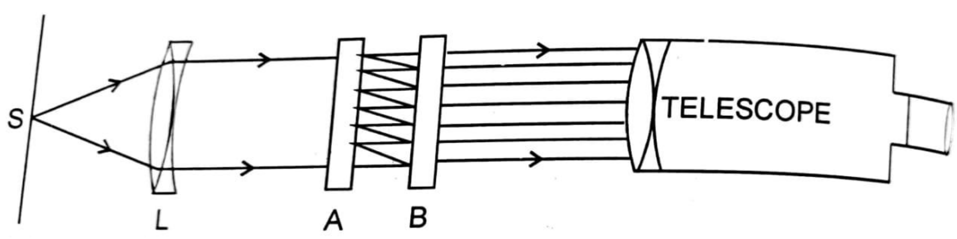


Figure : 18

8.12.4 Intensity Distribution

$$Reflectivity/Reflectance = R = \frac{intendity\ of\ reflected\ light\ ray}{intensity\ of\ incient\ light\ ray}$$

$$Transmittivily/Transmittance \ = T \ = \frac{intensity\ of\ transmitted\ light\ ray}{intensity\ of\ incident\ light\ ray}$$

We know Intensity $\alpha\ amplitude\ ^2 = \alpha \times \sqrt{I}$

$$\alpha_R\ \alpha\ \sqrt{I_R}\ \alpha\ \sqrt{R} = \ \alpha_R \Rightarrow \sqrt{R}\ \alpha$$

$$\alpha_T\ \alpha\ \sqrt{I_T}\ \ \alpha\ \sqrt{T} = \ \alpha_T \Rightarrow \sqrt{T}\ \alpha$$

Amplitudes of transmitted light rays are

a T, a T R, a T R^2, a T R^3, …….

Phase difference, $\delta = \frac{2\pi}{\lambda}\ (2d\ cos\theta)$

$y_1 = a\ T\ e^{i\omega t}$

$y_2\ = a\ T\ R\ e^{i(\omega t+\delta)}$

$y_3 = a\ T\ R^2 e^{i(\omega e+2\delta)}$

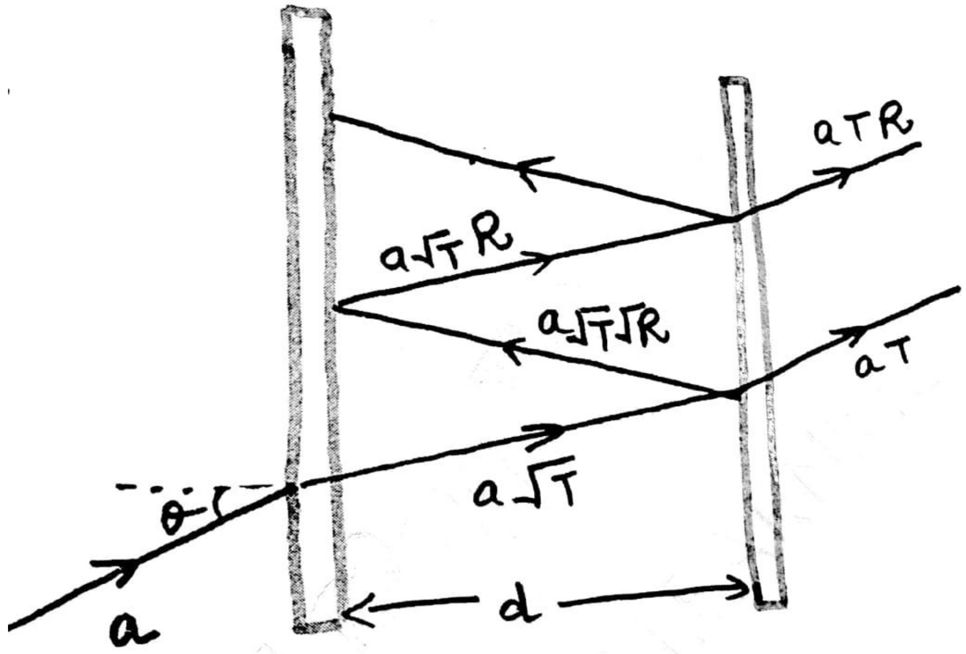


Figure : 19

The total/resultant amplitude (by using superposition principle).

$A = a\ T + a\ T\ R\ e^{i\delta} + a\ T\ R^2 e^{i2\delta} + \cdots \infty$

$$A = \frac{a\ T}{1 - R\ e^{i\delta}}$$

Intensity I = A = $\left(\frac{a\,T}{1-R\,e^{i\delta}}\right)\left(\frac{a\,T}{1-R\,e^{-i\delta}}\right) = \frac{a^2T^2}{1+R^2-R\left(e^{i\delta}+e^{-i\delta}\right)}$ (∵ $A^{*} complex\ conjugate\ of\ A$)

$$I = \frac{a^2T^2}{(1+R^2)-R\left(e^{i\delta}+e^{-i\delta}\right)}$$

$$= \frac{a^2T^2}{(1+R^2)-2R\,cos\delta}$$

we know

$$e^{i\delta} = cos\ \delta + i\ sin\ \delta$$

$$e^{ie\delta} = cos\ \delta - \ i^{o} sin\ \delta$$

$$\Rightarrow e^{i\delta} + e^{-i\delta} = 2\ cos\ \delta$$

$$\therefore cos\ \delta = 1 - 2\ sin^2\delta/2$$

$$1 - cos\ \delta = 2\ sin^{2\delta/2}$$

$$= \frac{a^2T^2}{(1+R^2-2R)+2R-2R\cos\delta}$$

$$= \frac{a^2T^2}{(1-R^2)+2R(1-\cos\delta)}$$

$$= \frac{a^2T^2}{(1-R)^2+4Rsin^{2\delta/2}}$$

$$I = \frac{a^2T^2/1-R^2}{\left[1+\frac{4R}{(1-R)^2}sin^2\delta/2\right]}$$

I will be max when,

$sin^2\delta/2$ = 0 ⇒ $sin\frac{\delta}{2} = 0$

$$\Rightarrow\ \frac{\delta}{2} = n\pi$$

$$\Rightarrow\ \delta = 2n\pi = 0, 2\pi, 4\pi\ ....$$

$$\Rightarrow\ I_{max} = \frac{a^2T^2/(1-R)^2}{1+0} \Rightarrow \boxed{I_{max} = \frac{a^2T^2}{(1+R)^2}}$$

So,

$$I = \frac{I_{max}}{1+\frac{4R}{(1+R)^2}}\ sin^2\delta/2$$

$$\boxed{I = \frac{I_{max}}{1+F\ sin^2\delta/2}}$$

Where, F is coefficient of Finesse $F = \frac{4R}{(1+R)^2}$

I will be minimum when $sin^2\delta/2 = 1$

$$I_{min} = \frac{I_{max}}{1+F}$$

$$\Rightarrow \frac{I_{max}}{I_{min}} = 1 + F$$

8.12.5 Visibility of Fringes

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = \frac{\frac{I_{max}}{I_{min}} - 1}{\frac{I_{max}}{I_{min}} + 1}$$

□ $\frac{I_{max}}{I_{min}} = 1 + F \quad \Rightarrow V = \frac{1+F-1}{1+F+1} = \frac{F}{F+2} = \frac{1}{1+\frac{2}{F}}$

□ $F = \frac{4R}{(1+R)^2} \Rightarrow V = \frac{1}{1+\frac{2\times(1-R)^2}{4R}} = \frac{2R}{2R+(1-R)^2}$

$$V = \frac{2R}{2R + 1 + R^2 - 2R}$$

$$\boxed{V = \frac{2R}{1 + R^2}}$$

$\Rightarrow \quad R = 0.5 \quad \Rightarrow \quad V = \frac{2\times 0.5}{1+0.25} = \frac{1}{1.25} = 0.8$

$\Rightarrow \quad R = 1(100\%) \Rightarrow V = \frac{2\times 1}{1+1} = 1$

with increase in reflectivity, V and hence sharpness increases i.e. contrast between the fringes becomes larger.

8.12.6 Sharpness of Fringes

It is determined by half width (where intensity is half of max. intensity)

$$I = \frac{I_{max}}{2}$$

$$\frac{I_{max}}{1 + F\, sin^2 \frac{\delta}{2}} = \frac{I_{max}}{2}$$

$$2 = 1 + FSin^2\delta/2$$

$$FSin^2\frac{\delta}{2} = 1 \quad \Rightarrow \; Sin^2\frac{\delta}{2} = \frac{1}{F}$$

$$Sin^2\frac{\delta}{2} = \frac{(1-R)^2}{4R} \;\Rightarrow\; Sin\;\frac{\delta}{2} = \frac{1-R}{2\sqrt{R}}$$

$$\frac{\delta}{2} = Sin^{-1}\left(\frac{1-R}{2\sqrt{R}}\right)$$

$\Rightarrow$ $\boxed{\delta = 2sin^{-1}\left(\frac{1-R}{2\sqrt{R}}\right)}$

It represents that the angular position from intensity maxima to the point where intensity becomes half of max. intensity.

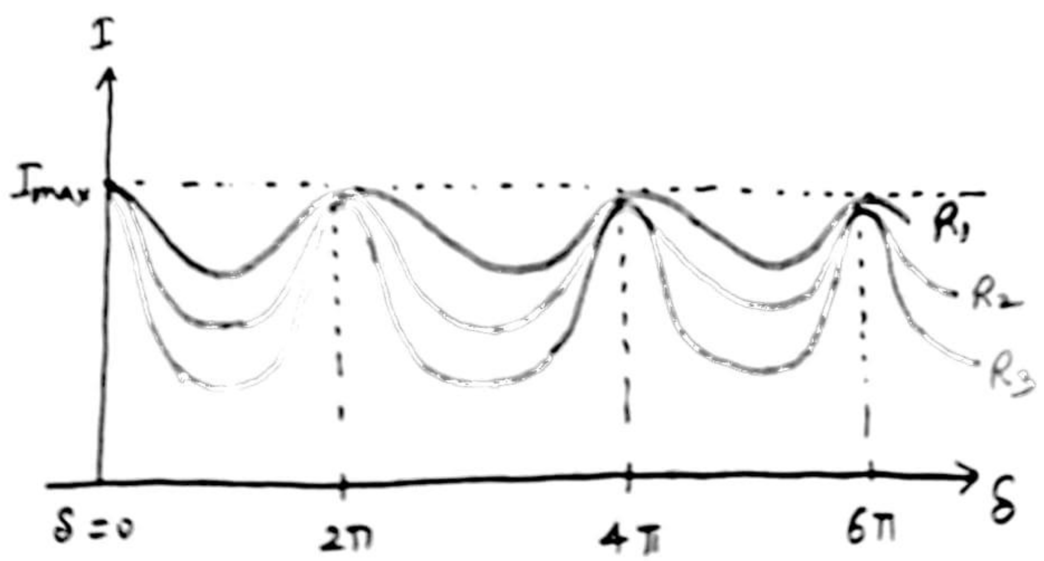


Figure : 20

8.12.7 Comparison with Michelson Interferometer

1. Febry-Perot fringes are much sharper than Michelson fringes.

Fabry-Perot

$\delta = 2\ sin^{-1}\left(\frac{1-R}{2\sqrt{R}}\right)$

for R = 0.8

$$\delta = 2\ sin^{-1}\left[\frac{1-0.8}{2\sqrt{0.8}}\right]$$

$$= 2\ sin^{-1}\left[\frac{0.2}{2\sqrt{\frac{4}{5}}}\right]$$

$$= 2\ sin^{-1}\left[\frac{1\sqrt{5}}{10\times 2}\right]$$

$$= 2\,sin^{-1}\left[\frac{1}{4\sqrt{5}}\right]$$

$$= 12.84^0 = 0.22\,rad$$

Michelson

$$I = I_{max}\,cos^2\frac{\delta}{2}$$

$$\frac{I_{max}}{2} = I_{max}\,cos^2\frac{\delta}{2}$$

$$\cos\frac{\delta}{2} = \frac{1}{\sqrt{2}}$$

$$\frac{\delta}{2} = 45^o$$

$$\delta = 90^o = \frac{\pi}{2} = 1.57\,rad$$

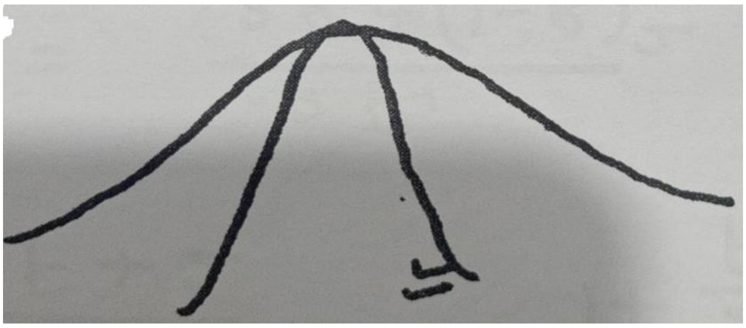


Figure : 21

half width in febry-perot is 7 times smaller than that in Michelson interferometer so sharpness is very large in febry-perot interferometer.

2. **Study of fine structure of spectral lines.**

If there are two or more wavelengths in the incident light (e.g. D_1 & D_2 of sodium light) then each wavelength produces its own fringe pattern in Febry-Perot interferometer. These fringes patterns are clearly separated from each other i.e. a good contrast between them. But in Michelson interferometer, the two wavelengths do not produce separate fringe patterns so their presence is checked by alternate distinctness and indistinctness of the fringes with increase in path difference due to increase in film thickness.

8.12.8 Applications of Febry-Perot Interferometer

1. **Measurement of λ of monochromatic light.**

$$\Delta x = 2d \cos\theta$$

At centre $\theta = 0^o$ $\Rightarrow \Delta x = 2d$

For bright fringes $\Delta x = 2d = n\lambda$

If the movable plate is displaced through a distance of $\frac{\lambda}{2}$ then a new fringe appears at centre.

Suppose by moving the movable plate through a distance Δd, the number of fringes appearing at centre are N.

$2(d + \Delta d) = (n + N)\lambda$ □ $\frac{\lambda}{2}$ .... = 1

$2d + 2\Delta d = n\lambda + N\lambda$ ∴ 1 .... = $\frac{1}{\lambda/2}$

$2\,\Delta d = N\lambda$ ∴ Δd .... = $\frac{\Delta d}{\lambda/2} = N$

$\boxed{\lambda = \frac{2\,\Delta d}{N}}$ $\frac{2\,\Delta d}{\lambda} = N$

2. **Spectral width/differences in wavelengths.**

Adjust the interferometer in such a position that bright fringes of both λ_1 & λ_2 coincide at centre

at plate separation d_1

$2d_1 = n_1\,\lambda_1 = n_2\,\lambda_2$

with increase the plate separation, suppose the bright fringes of λ_1 & λ_2 again coincide at plate separation d_1 for which m fringes of λ_1 and (m + 1) fringes of λ_2 appears.

$2\,d_2 = (n_1 + m)\,\lambda_1 \quad = \quad (n_2 + m + 1)\,\lambda_2$

$2\,d_2 = 2\,d_1 + m\,\lambda_1 \quad = \quad 2\,d_1 + (m + 1)\,\lambda_2$

$$\Rightarrow \qquad 2\,(d_2 - d_1) = m\lambda_1 \qquad \Rightarrow \qquad m = 2\frac{\Delta d}{\lambda_1} = 2\frac{\Delta d}{\lambda_2} - 1$$

$$\Rightarrow \qquad 2(d_2 - d_1) = (m+1)\,\lambda_2 \qquad \Rightarrow \quad 1 = 2\Delta d\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) = 2\,\Delta d\,\left(\frac{\lambda_1 - \lambda_2}{\lambda_1\,\lambda_2}\right)$$

$$(\lambda_1 - \lambda_2) = \Delta\lambda = \frac{\lambda_1\,\lambda_2}{2\,\Delta d} = \frac{\lambda^2}{2\,\Delta d} \qquad\qquad \text{where } d_2 - d_1 = \Delta d$$

8.13 L-G. PLATES

The **Lummer–Gehrcke interferometer** or **Lummer–Gehrcke plate** is a multiple-am interferometer similar to the Fabry–Pérot etalon, but using light at a steep angle of incidence. The interferometer consists of a long plate of glass or quartz, with faces that are polished accurately flat and parallel.[1] Light bounces back and forth inside the plate, striking the faces at an angle just below the critical angle as it propagates along.

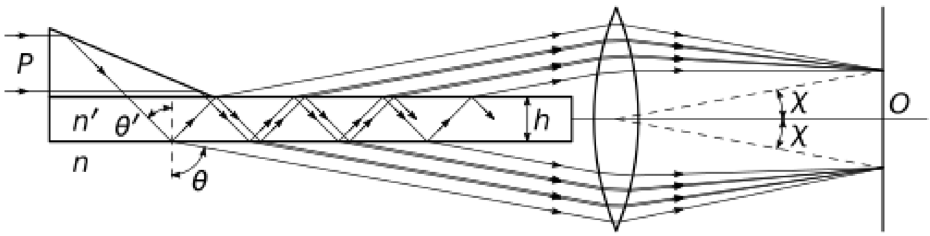


Figure : 22

Because of the steep angle of incidence, nearly all of the light is reflected, but a tiny fraction leaks out on each bounce. As in a Fabry–Pérot interferometer, the light that leaks out has phase that depends on how many times it has bounced inside the plate. A lens is used to overlap light that has emerged after varying numbers of bounces, producing an interference pattern. A key difference from a Fabry–Pérot etalon is that input light that reflects from the surface of the plate does not contribute to the interference.

Lummer–Gehrcke interferometers are now rarely used, having been largely replaced by Fabry–Pérot interferometers using modern dielectric reflective coatings.

8.14 SUMMARY

Interference by wedge-shaped film	■ The spacing β between two consecutive bright (or dark) fringes produced by wedge-shaped film is given by $$\beta = \frac{\lambda}{2\mu\theta}$$ where λ is the wavelength of light being used for illuminating the film, μ the refractive index of the film, and θ (measured in radians) the angle between the two plane surfaces, which form the wedge-shaped film.
	■ The diameters of the bright rings are proportional to the square-roots of the odd natural numbers, whereas the diameters of dark rings are proportional to the square-roots of natural numbers, provided the contact is perfect.
Newton's ring	■ On measuring the diameters of Newton's rings and the radius of curvature R, the wavelength can be calculated with the help of the following relation: $$\lambda = \frac{D_{n+p}^2 - D_n^2}{4pR}$$

8.15 TERMINAL QUESTIONS

1. Describe Michelson Interferometer with a neat diagram.
2. Give the theory of Newton's rings.
3. What are Newton's rings? Describe an experiment to determine the radius of curvature of a plane convex lens?
4. Describe the Principle and Working of Febry-Parot interferometer.
5. Write a short note on L-G Plates
6. Discuss application of Michelson's Interferometer.

8.16 SOLUTION & ANSWERS OF TERMINAL QUESTIONS

1. Section 8.11
2. Section 8.10
3. Section 8.10
4. Section 8.12, 8.121, 8.12.2
5. Section 8.13
6. Section 8.11.9

8-17 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.
2. Optics – Ajoy Ghatak
3. Optics – Principles and Application, K. K. Sharma, Academic Press.
4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.
5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT – 09: FRESNEL DIFFRACTION

STRUCTURE:

9.1 INTRODUCTION

The phenomenon of diffraction was first documented in 1665 by the Italian Francesco Maria Grimaldi. The use of lasers has only become common in the last few decades. The laser's ability to produce a narrow beam of coherent monochromatic radiation in the visible light range makes it ideal for use in diffraction experiments: the diffracted light forms a clear pattern that is easily measured. As light, or any wave, passes a barrier, the waveform is distorted at the boundary edge. If the wave passes through a gap, more obvious distortion can be seen. As the gap width approaches the wavelength of the wave, the distortion becomes even more obvious. This process is known as diffraction. If the diffracted light is projected onto a screen some distance away, then interference between the light waves create a distinctive pattern (the diffraction pattern) on the screen. The nature of the diffraction pattern depends on the nature of the gap (or mask) which diffracts the original light wave. Diffraction patterns can be calculated by from a function representing the mask. The symmetry of the pattern can reveal useful information on the symmetry of the mask. For a periodic object, the pattern is equivalent to the reciprocal lattice of the object. In conventional image formation, a lens focuses the diffracted waves into an image. Since the individual sections (spots) of the diffraction pattern each contain information, by forming an image from only particular parts of the diffraction pattern, the resulting image can be used to enhance particular features. This is used in bright and dark field imaging

When light from a narrow linear slit is incident on the sharp edge of an obstacle, it will be found that there is illumination to some extent within the geometrical shadow of the obstacle. This shows that light can bend round an obstacle. All phenomena like this which are produced when the incident wavefront is somehow limited are called diffraction of light. When waves encounter obstacles

(or openings), they bend round the edges of the obstacles, if the dimensions of the obstacles are comparable to the wavelength of the waves. The bending of waves around the edges of an obstacle is called diffraction.

In this unit, we shall study about types of diffraction and Fresnel construction of half period zones, propagation of light. Moreover, we also study about zone plates and cornu's spiral.

9.2 OBJECTIVES

After studying this unit, you should be able to –

- Describe the Fresnel diffraction.
- Explain the construction of half-period zones and compute their radii and area.
- Describe a zone plate, its construction, its action and theory.
- Difference between interference and diffraction.
- List the similarities and dissimilarities between a zone plate and convex tens.
- To understand the concept of diffraction at a straight edge.
- Understand the concept of cornu's spiral

9.3 WHAT IS DIFFRACTION

When light travels in air, it encounters various phenomena like interference, refraction, reflection and diffraction. When the light comes in contact with an obstacle, diffraction of light takes place.

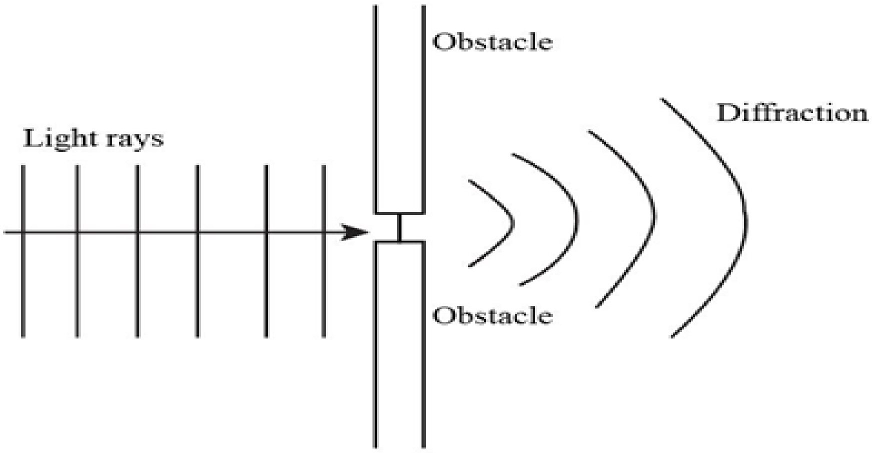


Figure : 1

When light passes through a small opening, comparable in size to the wavelength λ of the light, the wavefront on the other side of the opening resembles the wave. Let us know more about the diffraction of light and single slit diffraction that occurs when light travels through a single slit. Also, let us learn what happens in a single slit diffraction experiment.

We can observe single slit diffraction when light passes through a single slit whose width (w) is on the order of the wavelength of the light. The diffraction pattern on the screen will be at a distance L >> w away from the slit. The intensity is a function of angle

Figure : 2

The phenomenon of bending of light waves around edges of small obstacles and hence it's spreading into the geometrical shadow of the obstacle is called diffraction. The diffraction effects were first observed by **Grimaldi in 1665. The effects can be observed only when the size of the obstacle is very small and comparable to the wavelength of light.**

Figure : 3

This phenomenon shows that the rectilinear propagation of light (light traveling along a straight) is only approximate i.e., light bends at the corners of small obstacles and enters the regions of geometrical shadows. Light from the source S is made to fall on a slit AB whose width is very small. The region CD on the screen is found to contain unequally spaced alternate bright and dark fringes with light bending into the region above C and below D. This is due to diffraction effects. Fresnel explained this phenomenon by applying the Huygens ‘Principle along with the principle of interference.

9.3.1 Types of diffraction

There are two types of diffractions

- Fresnel Diffraction
- Fraunhofer Diffraction

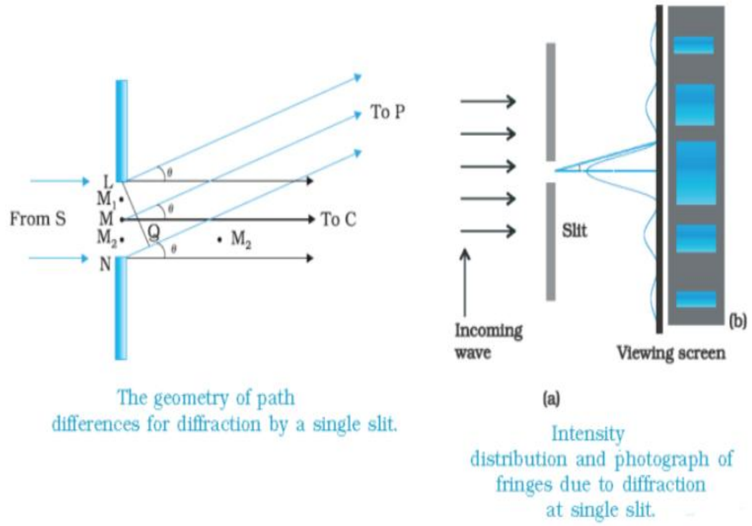


Figure : 4

From the above figure, we observe that the source is located at a finite distance from the slit, and the screen is also at a finite distance from the slit. The source and the screen are not very far from each other. So, this is a Fresnel diffraction. Here, if suppose the ray of light comes exactly at the edge of the obstacles, the path of the light is changed. So, the light bends a little and meets the screen.

A beam of width α travels a distance of α^2/λ, called the Fresnel distance before it starts to spread out due to diffraction. But when the source and the screen are far away from each other, and when the source is located at the infinite position, then the ray of light coming from that infinite source are parallel rays of light. So, this is Fraunhofer diffraction.

Here we have to make use of the lens. But why do we use the lens? Because in Fraunhofer diffraction, the source is at infinity so the rays of light which pass through the slit are parallel rays of light.

So, in order to make these rays parallel to focus on the screen, we, make use of the converging lens. The zone which we get in front of the slit is the central maxima. On either side of central maxima, there is bright zone i.e 1st maxima.

9.3.2 Fresnel Diffraction

Source, screen, and obstacle are close together.

The Fresnel diffraction is a process of diffraction that occurs when a wave passes through a slot and diffracts in the near field, causing any diffraction pattern observed to differ in size and shape, depending on the distance between the slot and the projection.

9.3.3 Fresnel's Assumptions

Fresnel in 1815, combined the Huygens principle of wavelet and the principle of interference to explain the bending of light around obstacles and also the rectilinear propagation of light.

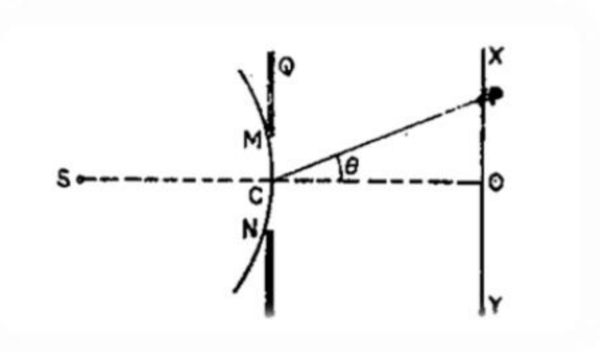


Figure : 5

1. According to Huygens' principle, each point of a wavefront (wavefront is a locus of points in a medium that are vibrating in same phase) is a source of secondary disturbance and wavelets coming from these points spread out in all directions with the speed of light. The envelope of these waves constitute the next wavelet.
2. According to Fresnel, a wavefront can be divided into a large number of strips or zones called Fresnel zones of small area. The resultant effect at any point will depend on the combined effect of all the secondary waves coming from various zones.
3. The effect at a point due to any particular zone depends on distance of the point from the zone.
4. The effect will also depend on the obliquity (inclination) of the point with reference to the zone under consideration.

9.3.4 Fraunhofer Diffraction

In Fraunhofer diffraction, the incident waves as well as the diffracted waves can be approximated to have plane wavefronts. This can be achieved by placing the source and the screen far away from the obstacle or by using lenses so that the source and the screen are at focal planes. Fraunhofer diffraction is also called far-field diffraction. Owing to the very nature of the wavefronts, Fraunhofer diffraction (planar wavefronts) is much simpler to analyse than Fresnel diffraction (spherical wavefronts). Therefore, we shall limit ourselves to Fraunhofer diffraction.

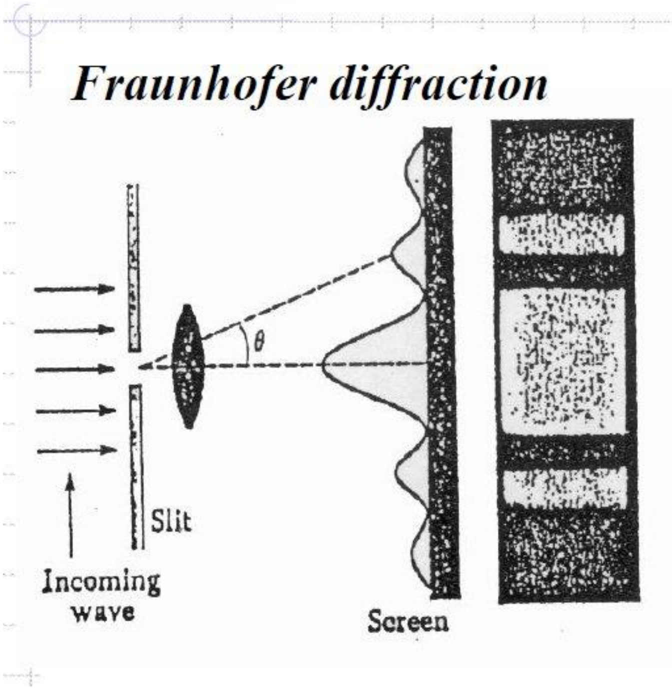


Figure : 6

9.3.5 Difference between Fresnel and Fraunhofer Diffraction

Fresnel's diffraction:	Fraunhofer's diffraction:
1. The source of light and the screen on which the diffraction pattern is observed are at **finite distance** from the obstacle or aperture. 2. The incident wavefront and the diffracted wavefronts are spherical or cylindrical. 3. The incident beam is a divergent beam whereas the diffracted beam is a convergent beam. 4. No changes in the wavefront are made by using either lenses or mirrors. 5. The centre of the diffraction pattern is either bright or dark. The pattern is the image of the obstacle or aperture.	1. The source of light and the screen on which the diffraction pattern is observed are at **infinite distance** from the obstacle or aperture. 2. The incident wavefront and the diffracted wave fronts are plane wave fronts. 3. The incident beam is a parallel beam and the diffracted beam is also parallel beam. 4. The incident rays from a source are made parallel using a convex lens and the diffracted rays are brought to focus on a screen using another convex lens (converging lenses). 5. The centre of the diffraction pattern is always bright. The pattern is the image of the source itself.

Examples of diffraction – (1) The luminous border that surrounds the profile of a mountain just before sun rises behind it, (2) the light streaks that one sees while looking at a strong source of light with half shut eyes and (3) the coloured spectra one sees while viewing a distant source of light through a fine piece of cloth.

9.4 DIFFERENCE BETWEEN INTERFERENCE AND DIFFRACTION

	Interference	Diffraction
1.	It is due to the superposition of secondary wavelets from two different wavefronts produced by two coherent sources.	It is due to the superposition of secondary wavelets emitted from various points of the same wave front.
2.	Frings are equally spaced.	Fringes are unequally spaced.
3.	Bright fringes are of same intensity.	Intensity falls rapidly.
4.	Comparing with diffraction, it has large number of fringes.	It has less number of fringes.

9.5 DIVISION OF WAVEFRONT INTO FRESMEL'S HALF PERIOD ZONE

Division of wavefront into Fresnel's half period zones – Expression for resultant displacement/amplitude – Rectilinear propagation of light

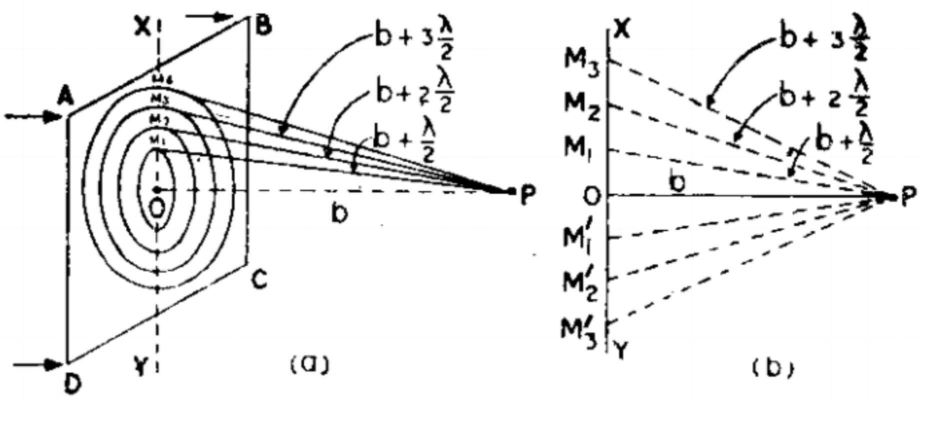


Figure : 7

ABCD is a plane wave front of monochromatic light of wavelength λ. The diagram shows the plane wavefront as perpendicular to plane of the paper. Consider a point P at a distance b from the wave front at which amplitude due to the wave is to be found.

To find the resultant amplitude at P due to entire wavefront, Fresnel assumed the wavefront to be divided into a number of concentric half period zones called

9.5.1 What is Fresnel's Half Period Zones

With P as centre and with $M_1P = \left(b + \frac{\lambda}{2}\right), M_1P = \left(b + \frac{2\lambda}{2}\right), \ldots \ldots$ as radii, a series of concentric spheres are drawn on the wavefront. These spheres intersect the wavefront in concentric circles. These circles or zones are of radii OM_1, OM_2,….. on the wavefront with O as centre.

The secondary waves from any two consecutive zones reach the point P with a path difference of $\frac{\lambda}{2}$ or a time period of $\frac{T}{2}$. Hence these zones are called half period zones. The area of the circle OM_1 is called first half period zone. The area between the circles of OM_2 and OM_1 is called second half period zone and so on. The area between the nth and $(n–1)^{th}$ circle is called the nth half period zone.

9.5.2 Radius of a Half Period Zone

In the diagram, from the right-angled triangle OM_1P,

$$OM_1 = \sqrt{(M_1P)^2 - (OP)^2} = \sqrt{\left(b + \frac{\lambda}{2}\right)^2 - b^2}$$

$OM_1 = \sqrt{\left(b^2 + 2b\frac{\lambda}{2} + \frac{\lambda^2}{4}\right) - b^2}$ or $OM_1 = \sqrt{b\lambda}$ (neglecting $\frac{\lambda^2}{4}$ as $b \gg \lambda$)

$OM_1 = \sqrt{b\lambda}$ is the radius of first half period zone.

The radius of the second half period zone is

$OM_2 = \sqrt{(M_2P)^2 - (OP)^2} = \sqrt{\left(b + \frac{2\lambda}{2}\right)^2 - b^2}$ Thus $OM_2 = \sqrt{2b\lambda}$

Similarly, the radius of the nth half period zone is $OM_n = \sqrt{\left(b + \frac{2\lambda}{2}\right)^2 - b^2}$

or $OM_n = \sqrt{nb\lambda}$

Thus, the radii of 1st, 2nd, …… half period zones are $\sqrt{b\lambda}, \sqrt{2b\lambda}, \dots \dots \sqrt{nb\lambda}$.

Therefore, the radii of the zones are proportional to the square root of natural numbers.

9.5.3 Area of Half Period Zone

The area of first half period zone is

$= \pi(OM_1)^2 = \pi[(M_1P)^2 - (OP)^2]$ (As area $= \pi r^2$)

$= \pi\left[\left(b + \frac{\lambda}{2}\right)^2 - b^2\right] = \pi\ b\ \lambda$

The area of 2nd half period zone $= \pi[(OM_2)^2 - (OM_1)^2]$

$$= \pi[2b\lambda - b\lambda] = \pi\ b\ \lambda$$

The are of nth half period zone $= \pi[(OM_n)^2 - (OM_1)^2]$

$$= \pi[nb\lambda - b\lambda] = \pi\ b\ \lambda$$

Thus, the area of each half period zone is same and is equal to λ .

Also, the area of any zone is directly proportional to the wavelength (λ) of light and the distance of the point from the wavefront (b).

9.5.4 Amplitude due to the wave front

The amplitude of the waves at P due to an individual zone is

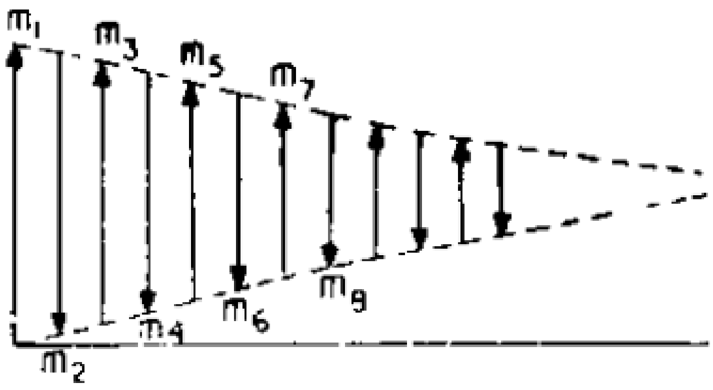


Figure : 8

1. Directly proportional to the area of the zone
2. inversely proportional to the distance of the point P from the given zone.
3. the obliquity factor (1+ cos θ) where θ is the angle between normal to the zone and the line *joining* the zone to the point P. The effect at P decreases as obliquity increases. The path difference between any two consecutive half period zones $\left(is\ \frac{\lambda}{2}\right)$. Hence the waves from two consecutive zones will reach P in opposite phase. If m_1, m_2, m_3,….. are the amplitudes at P due to 1st, 2nd, 3rd, ….. half period zones, the resultant amplitude at P due to entire wavefront is

   $A = m_1 - m_2 + m_3 - m_4 + \cdots . + m_n$ if n is odd

   and $A = m_1 - m_2 + m_3 - m_4 + \cdots . - m_n$ if n is even.

As the obliquity increases amplitudes decreases, ie. m_2 is less than m_1, m_3 is less than m_2 etc…

Thus on the average $m_2 = \frac{m_1+m_3}{2}$ …. (1)

Similarly $m_4 = \frac{m_3+m_5}{2}$. …….. (2)

The equation $A = m_1 - m_2 + m_3 - m_4 + \cdots$. can be written as

$$A = \frac{m_1}{2} + \left(\frac{m_1}{2} - m_2 + \frac{m_3}{2}\right) + \left(\frac{m_3}{2} - m_4 + \frac{m_5}{2}\right). \ldots\ldots \qquad \ldots\ldots\ldots.. (3)$$

Substituting equations (1) and (2) in (3) we get,

$A = \frac{m_1}{2} + \frac{m_n}{2}$ if n is odd …. (5) (The terms in the bracket cancel)

$A = \frac{m_1}{2} + \frac{m_{n-1}}{2} - m_n$ if n is even. …. (6).

As the amplitudes are of diminishing order, for large n, m_n and m_{n-1} tend to zero.

Thus, $A = \frac{m_1}{2}$.

The amplitude of the wave at any point P, in front of a large plane wavefront is equal to half the amplitude due to the first half period zone.

As the intensity is proportional to square of the amplitude, ($I \propto A^2$ the intensity at P is proportional to $\frac{m_1^2}{4}$ $\left(I \propto \frac{m_1^2}{4}\right)$. Thus, the intensity at point P is one fourth of the intensity due to the first half period zone.

9.6 RECTILINEAR PROPAGATION OF LIGHT

The intensity at point in front of a wave front is proportional to $\frac{m_1^2}{4}$, where m_1 is the amplitude of the first half period zone. Thus, the intensity at point P is one fourth of the intensity due to the first half period zone.

Thus, only half the area of the first half period zone is effective in producing the illumination at the point P. A small obstacle of the size of half the size of half the area of first half period zone placed at O will block the effect of whole wavefront and the intensity at P due to rest of the wavefront is zero.

While dealing with the rectilinear propagation of light, the size of the obstacle used is far greater than the area of the first half period zone and hence the bending effect of light round the corners of the obstacle is diffraction effects cannot be noticed.

Thus if the size of the obstacle placed in the path of light is very small and comparable to wavelength of light, then it is possible to observe illumination in the region of the geometrical shadow also. **Thus, rectilinear propagation of light is only approximately true.**

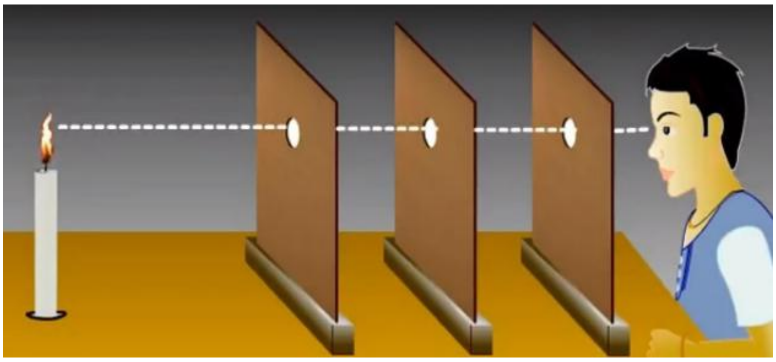


Figure : 9

9.7 WHAT IS ZONE PLATE

A zone plate is a specially constructed screen such that light is obstructed from every alternate zone. The correctness of Fresnel's method in dividing a wavefront into half period zones can be verified with the help of zone plate.

A zone plate is constructed by drawing concentric circles on a white paper such that radii are proportional to the square root of the natural numbers. The odd numbered zones (i.e. 1^{st}, 3^{rd}, 5^{th} …) are covered with black ink and a reduced photograph is taken.

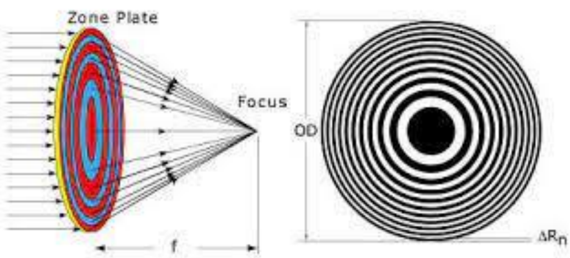


Figure : 10

9.7.1 Positive Zone Plate

The negative of the photograph appears is as shown in Fig. 11 (a). The negative shows odd zones are transparent to incident light and even zones will cut off light. This is a **positive zone plate**.

9.7.2 Negative Zone Plate

If odd zones are opaque and the even zones are transparent then it is a **negative zone plate**. Fig. 11(b)

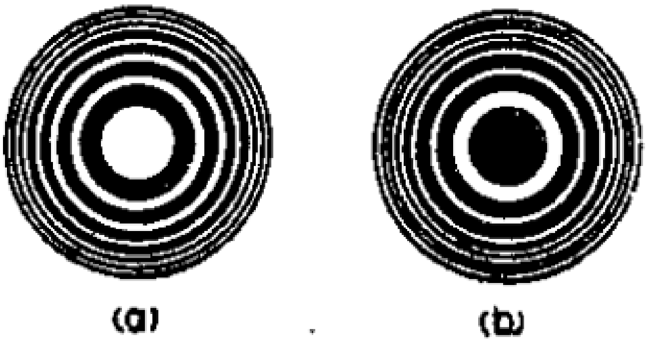


Figure : 11

9.7.3 Explanation of Zone Plate for an Incident Spherical Wave Front

Let S be a point source of light of wavelength λ placed at a distance a from centre O of the zone plate. Let P be the point on a screen placed at distance b at which intensity of diffracted light bright.

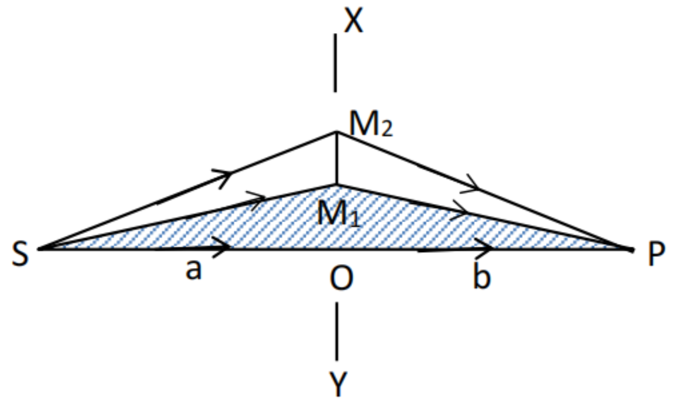


Figure : 12

Let r_1, r_2, r_3……. r_n be the radii of the 1st, 2nd, 3rd ……nth half period zones respectively. The position of the screen is such, that from one zone to the next there is an increasing path difference of $\frac{\lambda}{2}$.

Thus, from the diagram $SO + OP = a + b$

$SM_1 + M_1P = a + b + \frac{\lambda}{2}$ …….

(7)

Similarly, $SM_2 + M_2P = a + b + \frac{2\lambda}{2}$ and so on

From the triangle SM_1O $\quad SM_1 = (SO^2 + OM_1^2)^{1/2} = (a^2 + r_1^2)^{1/2}$

Similarly from the triangle PM_1O $\quad M_1P = (OP^2 + OM_1^2)^{1/2} = (b^2 + r_1^2)^{1/2}$

Substituting the values of SM_1 and M_1P in equation (7), we get

$(a^2 + r_1^2)^{1/2} + (b^2 + r_1^2)^{1/2} = a + b + \frac{\lambda}{2}$

or $a\left(1 + \frac{r_1^2}{a^2}\right)^{1/2} + b\left(1 + \frac{r_1^2}{b^2}\right)^{1/2} = a + b + \frac{\lambda}{2}$

Expanding and simplifying the above equation, we get

$a\left(1 + \frac{r_1^2}{2a^2}\right)^{1/2} + b\left(1 + \frac{r_1^2}{2b^2}\right)^{1/2} = a + b + \frac{\lambda}{2}$

$a + \frac{r_1^2}{2a^2} + b + \frac{r_1^2}{2b^2} = a + b + \frac{\lambda}{2}$

or $\frac{r_1^2}{2}\left(\frac{1}{a}+\frac{1}{b}\right)=\frac{\lambda}{2}$ $\qquad$ $or \quad r_1^2\left(\frac{1}{a}+\frac{1}{b}\right)=\lambda$

Thus, for the radius of the n^{th} zone the above relation can be written as

$$r_n^2\left(\frac{1}{a}+\frac{1}{b}\right)=n\,\lambda \ldots\ldots(8)$$

$or \quad r_n^2=\frac{ab}{a+b}\,n\,\lambda \qquad or \quad r_n=\sqrt{\frac{ab\lambda}{a+b}}\,\sqrt{n}$

Thus, the radii of the half period zones are proportional to the square root of the natural numbers.

From equation (8) can written as $\left(\frac{1}{a}+\frac{1}{b}\right)=\frac{n\lambda}{r_n^2}$ ………………… (9)

This equation is similar to the lens formula $\frac{1}{u}+\frac{1}{v}=\frac{1}{f}$ …………………. (10)

Comparing equations (9) and (10) $\frac{1}{f}=\frac{n\lambda}{r_n^2} \qquad f=\frac{r_n^2}{n\,\lambda}$

f is the focal length of zone plate and acts as a convex lens of multiple foci.

The path difference between any successive transparent zones is λ and the phase difference is 2π. Waves from successive zones reach P in phase.

9.7.4 Focusing action of Zone plate

The amplitude at P depends on (a) area of the zone, (b) distance of the zone from P and (c) obliquity factor.

The area of n^{th} zone $=\pi r_n^2-\pi\, r_{n-1}^2$

As $r_n^2=\frac{ab}{a+b}\,n\,\lambda$, the area of the n^{th} zone $\frac{ab}{a+b}\,n\,\lambda-\pi\frac{ab}{a+b}(n-1)\,\lambda=\pi\,\frac{ab\lambda}{a+b}$

Area is independent of n. Area of all zones are same. But the distance of the zone from P and obliquity factor increases as n increases.

The resultant amplitude at P is A = m1 + m3 + m5+…. for positive zone plate

A = - ($m_2 + m_4 + m_6$ +……) for negative zone plate.

This is much greater than $A=\frac{m_1}{2}$ which is due to all zones.

As the intensity from the zone plate is very high, the zone plate is said to have focussing action

9.7.5 Differences between Zone plate and Convex lens

Zone Plate	Convex Lens
• The Zone Plate has multiple focal length • Intensity of image is less • In Zone Plate the image is formed by diffraction of light. • The focal length of Zone Plate is $\frac{1}{f} = \frac{n\lambda}{r_n^2}$ • A Zone Plate is used over a wide range of Wavelength. • A Zone Plate has greater focal length for violet color than the red color $\boxed{f_v > f_r}$	• The convex lens has a single focal length. • Intensity of image is greater. • In a convex lens, the ray of light are focused to form an image by refraction. • The focal length of convex lens is $\frac{1}{f} = (n-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$ • A convex lens cannot be used over visible region. • A convex lens has greater focal length for red color than violet color $\boxed{f_r > f_v}$

9.8 EXPLANATION OF CYLINDRICAL HALF PERIOD STRIPS

S is a narrow rectangular slit or a linear source of light of wavelength λ. AB is the cylindrical wavefront. P is a point on the axis of the wavefront at which resultant intensity is to be found.

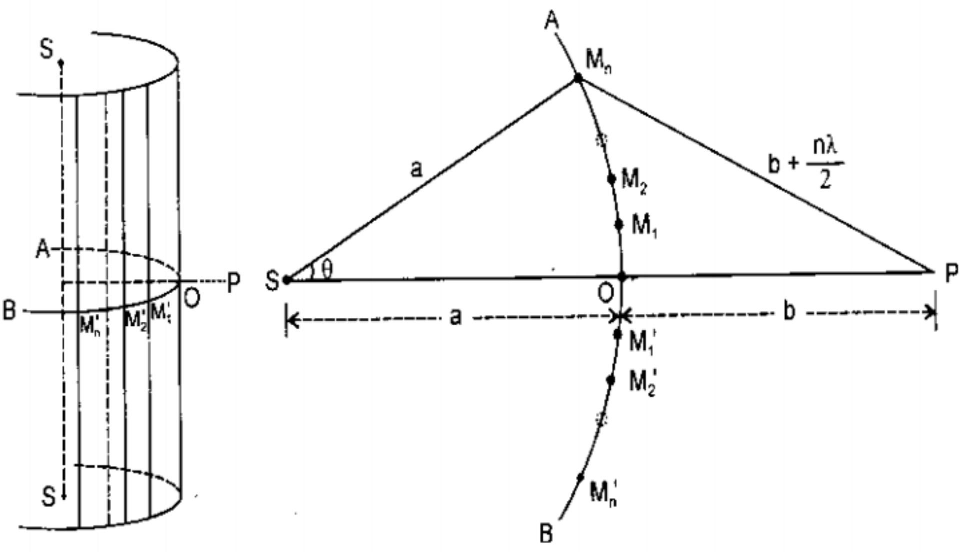


Figure : 13

To find the resultant amplitude/intensity at P due to the wavefront

If m_1, m_2, m_3 ……are the amplitudes at P due to 1st, 2^{nd}, 3^{rd}, …… half period strips on either side of O, the resultant amplitude due to one half of the wavefront is $A = m_1 - m_2 + m_3 - m_4 + ...$

$$A = \frac{m_1}{2} + \left(\frac{m_1}{2} - m_2 + \frac{m_3}{2}\right) + \left(\frac{m_3}{2} - m_4 + \frac{m_5}{2}\right) + \cdots$$

$$A = \frac{m_1}{2} \quad \text{(since } \frac{m_1+m_3}{2} = m_2\text{)}$$

Hence the resultant amplitude due to entire wavefront is $A = \frac{m_1}{2} + \frac{m_1}{2} = m_1$

9.8.1 Diffraction at a Straight Edge

Theory - S is a narrow rectangular slit illuminated light of wavelength λ. OE is the straight edge (opaque object covering half the slit in vertical plane), AB is the cylindrical wavefront, XY is the screen. CX is the region of diffraction fringes and CY has illumination of decreasing intensity.

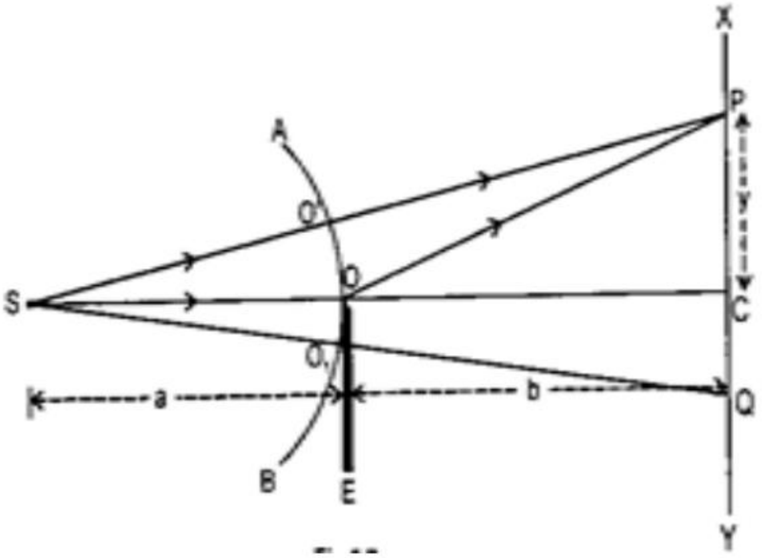


Figure : 14

The intensity at P at distance y from C on the screen is due to upper part of wavefront AB above O’ and that due to exposed portion OO’ of the wavefront.

The amplitude at P due to upper half above O’ is $= \frac{m_1}{2}$.

If portion OO’ contains one half period strip,

the total amplitude at $P = \frac{m_1}{2} + m_1 = \frac{3m_1}{2}$ (maximum amplitude)

If OO’ contains two half period strips,

then the total amplitude at $P = \frac{m_1}{2} + m_1 - m_2 = \frac{3m_1}{2} - m_2$ (minimum amplitude)

When region OO’ has odd half period strips, the amplitude is maximum and for even half period strips it is minimum. Above C, on the screen a diffraction pattern of alternate maxima and minima are observed. As distance on the screen increases, the intensity becomes uniform.

9.8.2 Condition for Maxima and Minima of Diffraction Pattern

The path difference between waves from O and O’ reaching P is $\delta = PO - PO'$

If δ is odd multiples of $\frac{\lambda}{2}$, amplitude at P is maximum.

$$\text{i.e. } \boldsymbol{\delta = PO - PO' = (2n+1)\frac{\lambda}{2}} \qquad \text{.......... (11)}$$

If δ is even multiples of $\frac{\lambda}{2}$, amplitude at P is minimum.

$$\text{i.e. } \boldsymbol{\delta = PO - PO' = 2n\frac{\lambda}{2} = n\lambda} \qquad \text{............ (12)}$$

where n = 0, 1, 2, 3, ……

From the diagram, $PO = \sqrt{(OC)^2 + (CP)^2}$

$$= \sqrt{b^2 + y^2} = b\left[1 + \frac{y^2}{b^2}\right]^{1/2} = b\left[1 + \frac{y^2}{2b^2}\right]$$

Thus $PO = b + \frac{y^2}{2b}$

and $PO' = SP - SO' = \sqrt{(SC)^2 + (CP)^2} - SO'$

$$= \sqrt{(a+b)^2 + y^2} - a$$

$$= (a+b)\left[1 + \frac{y^2}{(a+b)^2}\right]^{1/2} - a$$

Thus $PO' = (a+b)\left[1 + \frac{y^2}{2(a+b)^2}\right] - a = b + \frac{y^2}{2(a+b)}$

Hence $PO - PO' = b + \frac{y^2}{2b} - b - \frac{y^2}{2(a+b)}$

$$PO - PO' = \frac{ay^2 + by^2 - by^2}{2b(a+b)} = \frac{ay^2}{2b(a+b)} \qquad \text{........... (13)}$$

Comparing equation (11) with (13), the condition for maximum is

$\frac{ay^2}{2b(a+b)} = (2n+1)\frac{\lambda}{2}$ or $y^2 = \frac{2b(a+b)\,(2n+1)\lambda}{2a}$

or $\boldsymbol{y_n} = \sqrt{\boldsymbol{\frac{b(a+b)\,(2n+1)\lambda}{a}}}$. This is distance of n[th] maximum from the centre C.

Comparing equation (3) with (2), the condition for minimum is $\frac{ay^2}{2b(a+b)} = n\lambda$

or $\boldsymbol{y_n} = \sqrt{\boldsymbol{\frac{2b(a+b)n\lambda}{a}}}$. This is distance of n[th] minimum from the centre C.

The diagram shows the diffraction pattern due to straight edge.

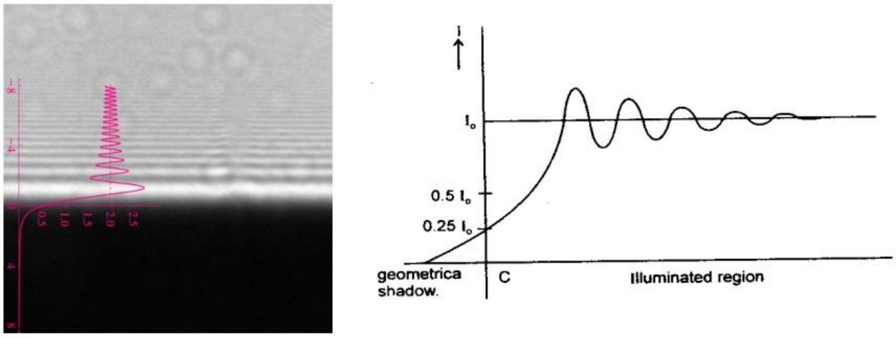


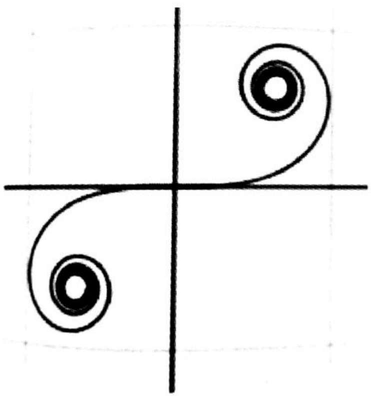
Figure : 15

The graph shows the intensity distribution due to diffraction at a straight edge. The intensity on the screen XY is due to upper part of wavefront AB only as the lower half is blocked.

The resultant amplitude at C is $\frac{m_1}{2}$ and the intensity at C is $\frac{{m_1}^2}{4}$. It is the one fourth of the intensity compared to intensity when entire wavefront is exposed.

9.9 CORNU'S SPIRAL

To find the effect at a point due to an incident wave front Fresnel's method consists in dividing the wavefront into half period strips or half period zones. The path difference between secondary waves from corresponding points of neighboring zones is equal to $\frac{\lambda}{2}$.



Cornu's Spiral

Figure : 16

In Fig. 17, S is a point source of light and XY is the incident spherical wave front. With reference to the point P, O is the pole of the wave front. Let a and b be the distances of the points S and P from the pole of the wave front. With P as centre and radius b, let us draw a sphere touching the incident wave front at O. The path difference between the waves travelling in the directions SAP and SOP is given by

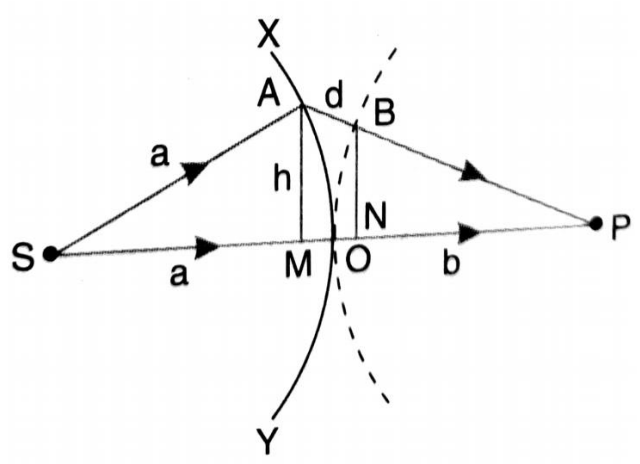


Figure : 17

$$d = SA + AP - SOP = SA + AP - (SO + OP) = a + AB + b - (a + b) = AB$$

$$d = AB$$

For large distances of a and b, AM and BN can be taken to be approximately equal and the path difference d can be written as

$d = AB = MO + ON$

But, from the property of a circle.

$$MO = \frac{AM^2}{2SO} = \frac{h^2}{2a} \text{ and } ON = \frac{BN^2}{2OP} = \frac{h^2}{2b} \qquad (approximately)$$

 $$d = \frac{h^2}{2a} + \frac{h^2}{2b} = \frac{h^2}{2ab}(a + b) \qquad \ldots\ldots\ldots\ldots\ldots\ldots (14)$$

If AM happens to be the radius of the nth half period zone, then this path difference is equal to $\frac{n\lambda}{2}$ according to the Fresnel's method of constructing the half period zones.

 $$\frac{h^2}{2ab}(a + b) = \frac{n\lambda}{2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

The resultant amplitude at an external point due to the wave front can be obtained by the following method. Let the first half period strip of the Fresnel's zones be divided into eight sub-strips and these vectors are represented from O to M_1 (Fig.18). The continuous phase change is due to the continuous increase in the obliquity factor from O to M_1. The resultant amplitude at the external point due to the first half period strip is given by OM_1 (= m_1). Similarly, if the process is continued, we obtain the vibration curve M_1M_2. The portion M_1M_2 corresponds to the second half period strip.

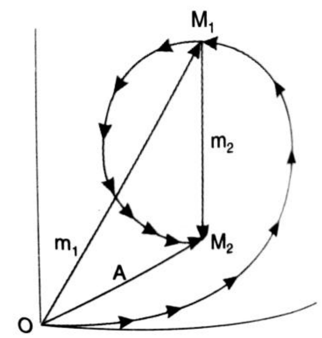


Figure : 18

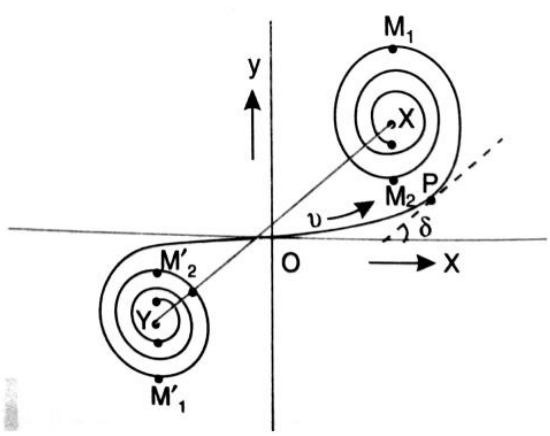


Figure : 19

The resultant amplitude at the point due to the first two half period strips is given by OM_2 (= A). If instead of eight sub-strips each period zone is divided into sub-strips of infinitesimal width, a smooth curve will be obtained. The complete vibration curve for whole wave front will be a spiral as shown in Fig. (18). X and Y correspond to the two extremities of the wave front and M_1 and M_2 etc. refer to the edge of the first, second, etc. half period strips. Similarly, M_1' and M_2' etc. refer to the edge of the first, second, etc. half period strips of the lower portion of the wave front. This is called

Carnu's spiral. The characteristic of this curve is that for any point P on the curve, the phase lag δ is directly proportional to the square of the distance ϑ. The distance is measured along the length of the curve from the point O. For a path difference of λ, the phase difference 2π. Hence, for a path difference of d, the phase difference δ is given by,

$$\delta = \frac{2\pi}{\lambda} d.$$

Substituting the value of d from equation (15) we get,

$$\delta = \frac{\pi}{2}\left[\frac{2h^2(a+b)}{ab\lambda}\right] \qquad \text{……… (16)}$$

$$\delta = \frac{\pi}{2}\vartheta^2 \qquad \text{……… (17)}$$

The value of ϑ is given by

$$\vartheta^2 = \frac{2h^2(a+b)}{ab\lambda}$$

or $$\vartheta = h\sqrt{\frac{2(a+b)}{ab\lambda}} \qquad \text{……… (18)}$$

Cornu's spiral can be used for any diffraction problem irrespective of the value of a, b and λ.

9.10 SUMMARY

In the present unit, we have studied about basic idea about diffraction phenomena together with various types of diffraction have been discussed. Diffraction pattern is formed whenever a wave encounters an object or aperture, the size of which is comparable to wavelength of light. To make the concept more clear the difference between interference and diffraction, construction and theory of Fresnel's half period zones and zone plate are explained.

We have also discussed focusing action of Zone Plate and diffraction at a straight edge.

9.11 TERMINAL QUESTIONS

1. Define Diffraction Phenomenon of light what are its type?
2. What is a zone plate? How does if differ from a convex tens.
3. Differentiate between interference and diffraction of light.
4. What are Fresnel half period zones?
5. Explain in brief rectilinear propagation of light.
6. Differentiate between Fresnel and Fraunholfer diffraction.
7. Explain the construction of Fresnel's half period zones on a plane wave front.
8. Discuss the Properties of Cornu's spiral and explain its relationship with Fresnel's half period zones.
9. Describe Focusing action of a zone plate.
10. Give the theory of diffraction at a straight edge and show that the Bands produced are not equally spaced.

9.12 ANSWER AND SOLUTIONS OF TERMINAL QUESTION

1. Section 9.3
2. Section 9.7, 9.7.5
3. Section 9.4
4. Section 9.5.1
5. Section 9.6
6. Section 9.3.5
7. Section 9.5
8. Section 9.9
9. Section 9.7.4
10. Section 9.8.1

9.13 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.
2. Optics – Ajoy Ghatak
3. Optics – Principles and Application, K. K. Sharma, Academic Press.
4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.
5. Advanced Geometrical Optics – Psang Dain Lin.

UNIT – 10: FRAUNHOFER DIFFRACTION

STRUCTURE:

10.1 INTRODUCTION

In this Unit, we shall study the Fraunhofer class of diffraction phenomenon also diffraction effects produced by diffracting elements when the source and the screen are effectively infinite distance from them when the source and the screen are at infinity, the wavefront incident on the diffracting element is plane. This condition can also be realized in practice by using a collimating lens, keeping the slit source at a distance equal to the focal length of the lens. Similarly, the diffraction effects are observed by focusing the diffracted rays by a convex lens.

In the Present Unit, we also come across single and double slit Fraunhofer diffraction and last, we will discuss grating and prism spectra.

10.2 OBJECTIVES

After studying this unit, you should be able to –

- Define Fraunhofer Diffraction
- Explain the Single and Double Slit Fraunhofer Diffraction.
- Introduce the concept of Plane diffraction grating.
- Have the basic idea of diffraction and its various classes.
- Describe the concept of Rayleigh criterion of resolution.
- Calculate the resolving power of various & instrument &/accessories like grating, prism, telescope and microscope.

10.3 WHAT IS FRAUNHOFER DIFFRACTION

Fraunhofer diffraction is **the type of diffraction that occurs in the limit of small Fresnel number**. In Fraunhofer diffraction, the diffraction pattern is independent of the distance to the screen, depending only on the angles to the screen from the aperture.

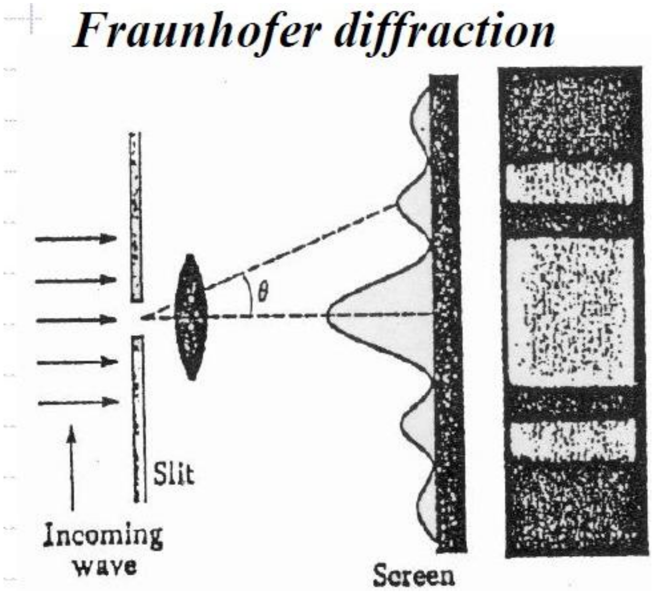


Figure : 1

10.4 THEORY OF SINGLE SLIT FRAUNOLFER DIFFRACTION

In FRAUNHOFER DIFFRACTION the source and the screen are an infinite distance from the obstacle and the wavefront is plane.

Let a parallel beam of monochromatic light of wavelength be incident normally upon a narrow-slit AB of width e where it gets diffraction in below fig. if a lens L is placed in the path of the diffraction beam, a real image of the diffraction pattern is formed on the screen MN in focal plane of the lens.

$$BE = AB \sin\theta = e \sin\theta$$

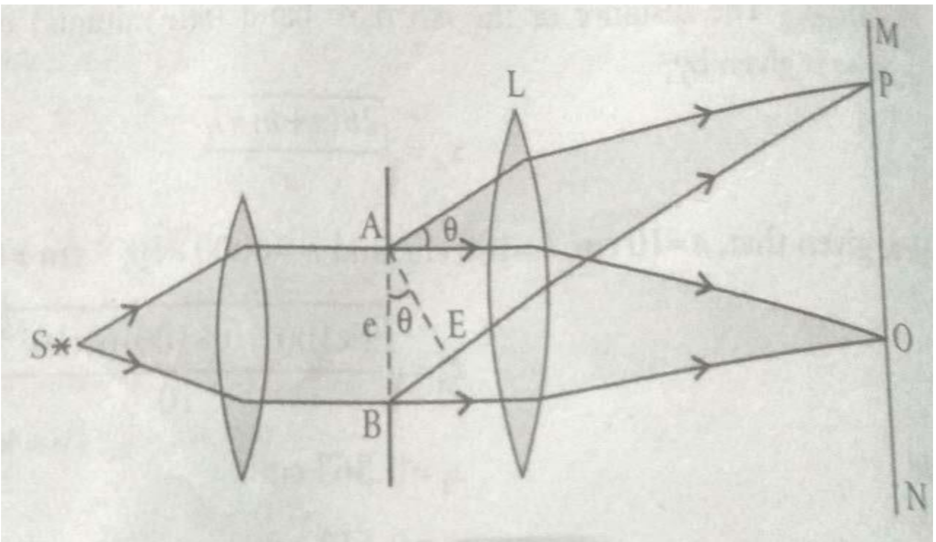


Figure : 2

The corresponding phase difference = $\left(\frac{2\pi}{\lambda}\right)^*$ path difference

$$= \left(\frac{2\pi}{\lambda}\right)^* (\text{e } \sin\theta)$$

Now, consider the width AB of the slit divided into n equal parts. Each part forms an elementary source. The amplitude of vibration at P due to the wave from each part will be the same, and the phase difference the waves from any two consecutive parts is,

$$= \frac{1}{n}\left(\frac{2\pi}{\lambda} e \, sin\theta\right) = d \; (let)$$

Hence, resultant amplitude at P is given by,

$$A = \frac{a \sin\left(\frac{nd}{2}\right)}{\sin\left(\frac{d}{2}\right)}$$

$$= \frac{\text{asin}\left(\frac{\pi esin\theta}{\lambda}\right)}{\sin\left(\frac{\pi \; esin\theta}{n\lambda}\right)}$$

Let $\frac{\pi e \; sin \; \theta}{\lambda} = \alpha, then$

$$A = \frac{asin\alpha}{sin\left(\frac{\alpha}{n}\right)} = \frac{asin\alpha}{\alpha/n} = \frac{na \; sin\alpha}{\alpha}$$

$$\left(\because \frac{\alpha}{n} \text{ is very small}\right)$$

If $n \to \infty, \alpha \to 0$, but the product na remains finite. Let na = A_0, then

$$A = \frac{A_0 \, sin\alpha}{\alpha}$$

So the resultant intensity at P will be,

$$I = A^2 or\ I = \left(\frac{A_0 \ sin\alpha}{\alpha}\right)^2 \quad \text{…………..} \ (1)$$

Since the magnitude intensity at any point in the focal plane of the lens is a function of α and θ so, we obtain a series of maxima and minima

Condition For Minima

From eq. (1) it is clear that the intensity is zero, wen

$$\text{Sin}\ \alpha = 0$$

$$\left[but\ \alpha \neq 0, \frac{sin\alpha}{\alpha} = 1, when\ \alpha = 0\right]$$

or $\alpha = \pm n\pi$

Where n is the integer except zero.

But, we know that,

$$\alpha = \frac{\pi esin\theta}{\lambda}$$

So, the position of minima are, given by

$$\frac{\pi esin\theta}{\lambda} = \pm n\pi$$

Or $esin\theta = \pm n, = \pm\lambda, \pm 2\lambda, \pm 3\lambda, and\ so\ on$

The eq. given the directions of first, second, third, ………… and so on minima.

Condition For Maxima

To locate the position of maxima of intensity in the diffracting pattern, let us differentiate I with respect to α and equal to zero i.e.,

$$\frac{dl}{d\alpha} = 0$$

$$\frac{d\left(\frac{A_0\ \ sin\alpha}{\alpha}\right)^2}{d\alpha} = 0$$

$$A_0^2\left[\frac{\alpha^2 2\sin\alpha\cos\alpha - \ sin^2\alpha.2\alpha}{\alpha^4}\right] = 0$$

$$\alpha^2\ 2\sin\alpha\cos\alpha - \ sin^2\alpha.\alpha = 0$$

$$\alpha\sin\alpha\,[\alpha\cos\alpha - \sin\alpha] = 0$$

$$\alpha\sin\alpha = 0$$

$$\alpha\cos\alpha = [\sin\alpha] = 0$$

$$\alpha = \tan\alpha$$

This eq. is solved graphically by plotting the

Y = α ……………….. (2)

Y = tan α ………………. (3)

The point of intersection of these two curves given the roots of eq. α = tan α Eq. (2) and (3) are shown in graph.

$$\boldsymbol{\alpha = 0, \frac{3\pi}{2}, \frac{5\pi}{2}, \frac{7\pi\ldots}{2}\ldots\ldots}$$

Or in general,

Substituting these value of in the equation (1), we get

$$I_{0_{\alpha\to 0}} = (\frac{A0\ sin0}{0})^2 = {A_0}^2$$

$$I_{0_{\alpha\to \frac{3\pi}{2}}} = (\frac{A0\ sin\frac{3\pi}{2}}{\frac{3\pi}{2}})^2 = \frac{4}{9\pi^2}{A_0}^2 = \frac{{A_0}^2}{22}$$

$$I_{0_{\alpha\to \frac{5\pi}{2}}} = (\frac{A0\ sin\frac{5\pi}{2}}{\frac{5\pi}{2}})^2 = \frac{4}{25\pi^2}{A_0}^2 = \frac{{A_0}^2}{61}$$

Thus, the intensity of the successive maxima are in the ratio

$$1 : \frac{1}{22} : \frac{1}{61} : \frac{1}{121} \ldots\ldots.$$

Cleary, most of the incident light is concentrated in the principal maxima which occurs in the direction given by α = 0

$$\frac{\pi e sin\alpha}{\lambda} = 0 \qquad \text{and} \qquad \theta = 0$$

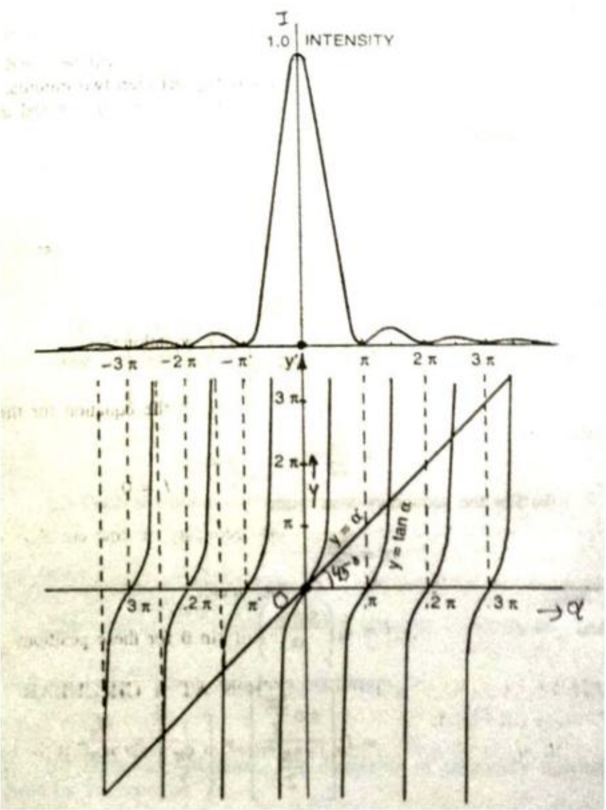


Figure : 3

i.e., in the direction of the incident light. Thus, the diffraction pattern consist of direction of the incident light, having maxima on either side of it, as shown in above fig. 3 the maxima lie at $\boldsymbol{\alpha\alpha}$ = ± $\boldsymbol{\pi\pi}$, ± $\mathbf{2222}$, ± $\mathbf{3333}$ … … …

The weak maxima do not fall exactly mid way between two minima, but are displaced towards the centre of the pattern by a certain amount which decreases with increasing order.

10.5 THEORY OF DOUBLE SLIT FRAUNHOLFER DIFFRACTION

Let a plane wavefront be incident normally on slit S_1 and S_2 equal e and separated by an opaque distance d. The diffracted light is focused on the screen XY. The diffracted pattern on the screen consists of equally spaced bright and dark fringe due to interference of light from both the slits and modulated by diffraction pattern from individual slits.

The diffraction pattern due to double-slit can be explained considering the following points :

- All the points in slits S_1 and S_2 will send secondary waves in all directions.
- All the secondary waves moving along the incident wave will be focused at P and the diffracted waves will be focused at P′.
- The amplitude at P′ is the resultant from two slit each of amplitude $R = \frac{A\, sin\alpha}{\alpha}$.
- T two waves from two-slit S_1 and S_2 will interfere at P′.

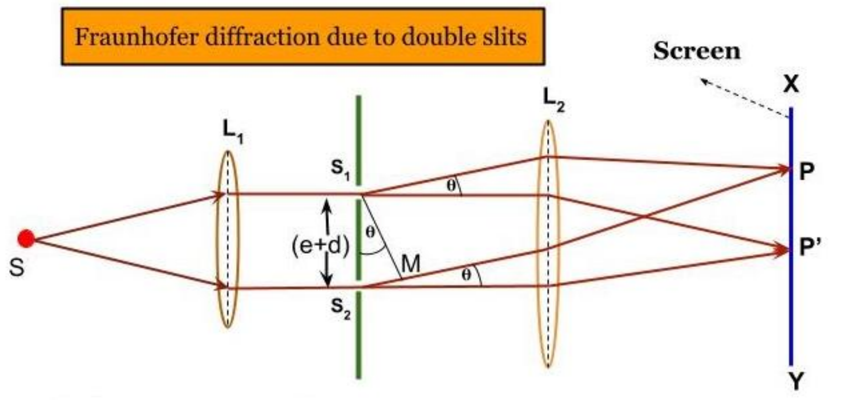


Figure : 4

$$\Delta = S_2M$$

$$\Delta = (e+d)sin\theta \qquad \text{……………} (4)$$

The corresponding phase difference :

$$\Delta\phi = \frac{2\pi}{\lambda}(e+d)sin\theta \qquad \text{………….. } (5)$$

$$\text{Let } \Delta\phi = 2\beta \qquad \text{……………. } (6)$$

$$\beta = \frac{\pi}{\lambda}(e+d)sin\theta \qquad \text{…………. } (7)$$

The resultant amplitude at P′ can be obtained by the vector addition method. The resultant amplitude at P′.

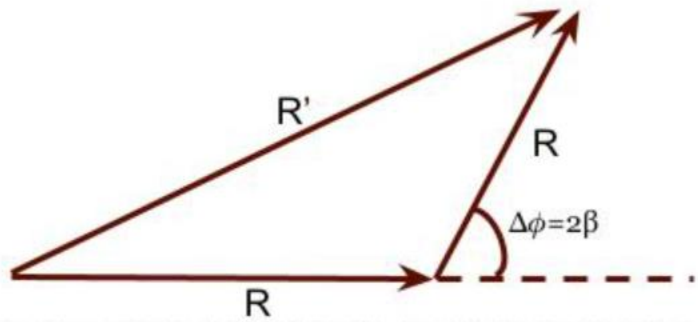


Figure : 5

$$R'^2 = R^2 + R^2 + 2R.\,R.\,cos\Delta\phi$$

$$R'^2 = R^2 + R^2 + 2R.\,R.\,cos2\beta \text{ \{\textbackslash qquad } \because 2\beta = \Delta\phi\}$$

$$R'^2 = 2R^2 + 2R^2 cos2\beta$$

$$R'^2 = 2R^2\,(1 + cos2\beta)$$

$$R'^2 = 4R^2 cos^2\beta \qquad \text{………… (8)}$$

Where R – Resultant amplitude of each slit S_1.

$$R = \frac{A\ sin\alpha}{\alpha} \qquad \text{………….. (9)}$$

Substituting the value of R in equation (8)

$$R'^2 = 4A^2\,\frac{sin^2\alpha}{\alpha^2} cos^2\beta \qquad \text{…………. (10)}$$

$$\boxed{R' = 2A\frac{sin\alpha}{\alpha}cos\beta} \qquad \text{……… (11)}$$

The intensity of the resultant diffraction pattern at P′

$$\boxed{I = 4A^2\frac{sin^2\alpha}{\alpha^2}cos^2\beta} \qquad \text{…………. (12)}$$

Where $\alpha = \frac{\pi}{\lambda} e\ sin\theta$ …………… (13)

The resultant intensity at any point is the contribution of the following two factors :

- The factor $\frac{A^2 sin^2\alpha}{\alpha^2}$, represents the intensity distribution due to diffraction from any individual slits.
- The factor $\cos^2\beta$ represents the distribution due to interference of waves from two parallel slits.

1. Maxima and minima due to diffraction term

2. **Maxima and minima due to interference term**

1. **Maxima and minima due to diffraction term :**

(i) **Principal Maxima :**

The diffraction term $\frac{A^2 sin^2\alpha}{\alpha^2}$ gives the central maxima, for α = 0 so

$$\boxed{\frac{\pi}{\lambda} e\ sin\theta = 0}$$

$$sin\theta = 0$$

$$\boxed{\theta = 0}$$

(ii) **Minima :**

The diffraction term $\frac{A^2 sin^2\alpha}{\alpha^2}$ gives the central maxima, for *sin*α = 0 so

$$\alpha = \pm m\pi$$

$$e\ sin\theta = \pm m\pi$$

(iii) **Secondary Maxima :**

The secondary maxima are obtained in the direction given by :

$$\boxed{\alpha = \pm\frac{3\pi}{2}, \pm\frac{5\pi}{2}, \pm\frac{7\pi}{2}, \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots}$$

2. **Maxima and minima due to interference term :**

(i) **Maxima :**

The interference term $\cos^2\beta$ gives maxima in the direction :

$$cos^2 = 1$$

$$\beta = \pm n\pi$$

$$\frac{\pi}{\lambda}(e+d)sin\theta = \pm n\pi$$

$$(e+d)sin\theta = \pm n\lambda \qquad \ldots\ldots\ldots\ldots (14)$$

Where n = 0, 1, 2, 3, …….

In the direction $\theta = 0^o$, the principle maxima due to interference and diffraction coincide.

(ii) Minima :

The interference term $\cos^2 \beta$ gives minima in the direction :

$$cos^2\beta = 0$$

$$\beta = \pm(2n+1)\frac{\pi}{2}$$

$$\boxed{(e+d)\,sin\theta = \pm(2n+1)\frac{\pi}{2}} \qquad \text{……….. (15)}$$

The intensity distribution curve due to the diffraction term, interference term, and the combined effect is shown in the figure below :

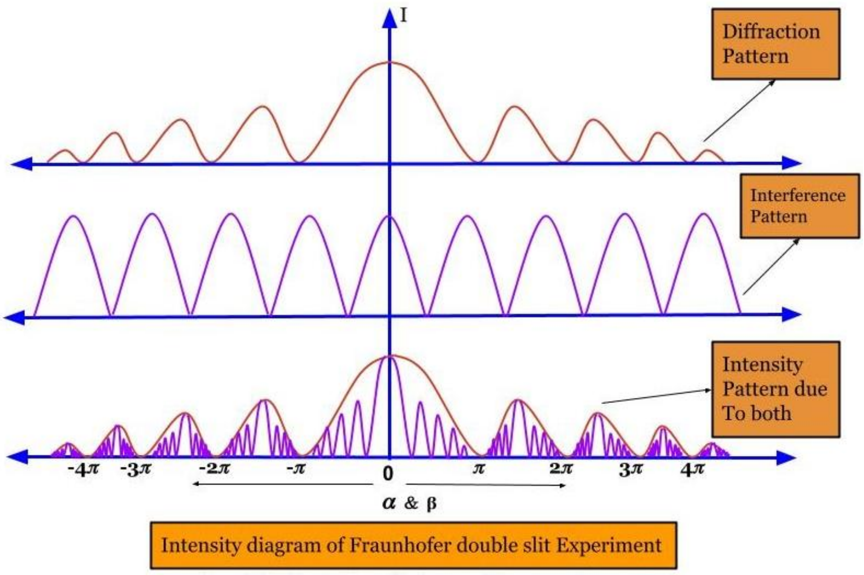


Figure : 6

10.5.1 Importance Fraunhofer Diffraction

In this type of diffraction, the source of light and the screen are at an effective infinite distance from the diffracting object (i.e., the slit). For this reason, it is called far-field diffraction. Since the source is at an infinite distance from the slit, the incident wavefront is usually a plane wavefront. The diffraction pattern remains the same irrespective of any shift in the object (slit) position. **The shape and intensity of the diffracting pattern remains the same throughout.**

10.6 PRINCIPAL OF PLANE DIFFRACTION GRATING

A diffraction grating is an optical component with a usual model. The form of the light diffracted by a grating depends on the construction of the elements. It is an optical component whose effect is alike to a prism: it splits white light into its component colors.

A plane transmission grating is a transparent plate or surface made of glass or of similar material on which a very large number of equidistant parallel lines very near to each other are scratched by a sharp diamond point. Here each line acts as an opaque line and the space between two lines acts as a thin slit. Each centimeter of this grating contains 5000 to 6000 lines. It can be easily constructed by drawing a huge number of intimately spaced lines on a plane transparent plate-like glass with a pointed diamond point. The lines on the plate are opaque to light and the spaces between these lines are transparent.

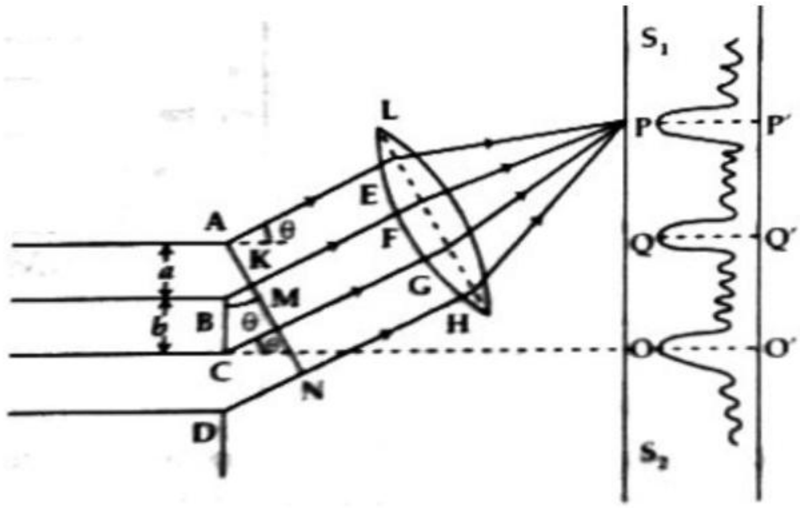


Figure : 7

For ordinary use, another type of transmission grating is used in the laboratories. Replicas of the original grating are prepared on a celluloid film.

On the original grating surface, a thin layer of collodion solution is poured and the solution is allowed to harden. Then the film of collodion is removed from the grating surface and then fixed between two glass plates. This serves as a plane transmission grating. It is called replica grating.

10.6.1 Theory of Diffraction Grating (Normal Incidence)

Consider parallel beam of light striking the transmission diffraction grating MN. The waves from different slits superpose and produce diffraction pattern on the screen. The pattern consists of a number of principal maxima with minima and secondary maxima in between. The incident beam travelling in the same direction will be brought to focus at O which corresponds to central maximum.

To find the intensity at P_1 - Fraunhofer diffraction at a single slit is applied.

The wavelet travelling from all the points in a slit along the direction θ are equivalent to a single wave of amplitude R $= A\frac{sin\alpha}{\alpha}$ where $\alpha = \frac{\pi d sin\theta}{\lambda}$ t

If there are N slits, there are N waves each from middle of the slits.

The path difference between any two consecutive slits is

$\delta = CG = AC\, sin\theta = (a+b)sin\theta$

[From diagram above, consider the triangle ACG, where CG is the path difference and $sin\theta = CG/AC$ where AC = (a + b)

The phase difference $= \frac{2\pi}{\lambda} \times (a + b)\, \sin\theta$

This is a constant and let it be equal to 2β

$2\beta = \frac{2\pi}{\lambda} \times (a + b)\, \sin\theta$ or $\beta = \frac{\pi(a + b)\, \sin\theta}{\lambda}$

By the method of vector addition of amplitudes, the resultant amplitude in the direction of θ is R $= A\frac{sin\alpha}{\alpha}\ \frac{sinN\beta}{sin\beta}$

[By vector addition $R = a\,\frac{sin\frac{n\phi}{2}}{sin\frac{\phi}{2}}$ and here a $= A\frac{sin\alpha}{\alpha}$, $n = N$ and $\phi = 2\beta$

The resultant intensity $I = R^2 = \left(A\frac{sin\alpha}{\alpha}\right)^2\left(\frac{sinN\beta}{sin\beta}\right)^2$

The factor $\left(A\frac{sin\alpha}{\alpha}\right)^2$ gives distribution of intensity due to single slit and the factor $\left(\frac{sinN\beta}{sin\beta}\right)^2$ gives distribution of intensity as combined effects of all the slits.

Condition for principal maxima

The intensity would be maximum when $sin\beta = 0$,

or $\beta = \pm n\pi$ where n = 0, 1, 2, …

At the same time $sinN\beta = 0$, so that the factor $sinN\beta/sin\beta$ becomes indeterminate.

It is evaluated as follows

$$\lim_{\beta\to\pm n\pi}\frac{sinN\beta}{sin\beta} = \lim_{\beta\to\pm n\pi}\frac{\frac{d}{d\beta}(sinN\beta)}{\frac{d}{d\beta}(sin\beta)} = \lim_{\beta\to\pm n\pi}\frac{NcosN\beta}{cos\beta} = \pm N$$

Hence $\lim_{\beta\to\pm n\pi}\left(\frac{sinN\beta}{sin\beta}\right)^2 = N^2$

The resultant intensity is $I = R^2 = \left(A\frac{sin\alpha}{\alpha}\right)^2 N^2$

The maxima are referred to as **principal maxima**.

The maxima are obtained for $\beta = \pm n\pi$ or $\beta = \frac{\pi(a+b)\sin\theta}{\lambda} = \pm n\pi$

or $\boldsymbol{(a + b)\, sin\theta = \pm n\lambda}$ where *n = 0, 1, 2, 3,…….*

n = 0 → central maximum

n = 1, 2, 3,…. → first, second, third,… … principal maxima.

Condition for minima

A number of minima occur, when $sinN\beta = 0$ but $n\beta \neq 0$.

Thus $sinN\beta = 0$ implies $N\beta = \pm m\pi$

$$N\beta = N\frac{\pi(a+b)\sin\theta}{\lambda} = \pm m\pi$$

or $N(a + b)\sin\theta = \pm m\lambda$ where m has all integral values except *0. N, 2N,…….nN,* since for these values $sin\beta = 0$ corresponding to principal maxima. Thus *m = 1, 2, 3, ……(N – 1).*

Condition for secondary maxima

As there are (N – 1) minima between two adjacent principal maxima, there must be (N – 2) other maxima between two principal maxima. These are known as secondary maxima. As N becomes large, intensity of these maxima decreases relative to principal maxima and become negligible.

The graph shows the intensity distribution curve under different conditions as shown.

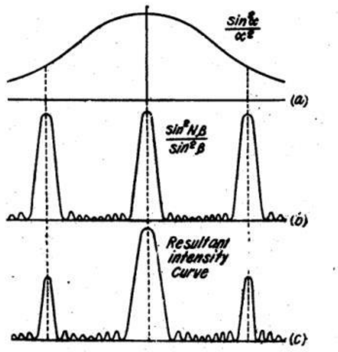


Figure : 8

Width of principal maxima in the diffraction pattern

The condition for the n^{th} maxima in the diffraction pattern due to grating is given by

$$(a + b) \sin\theta_n = n\lambda \qquad \text{……………..} (16)$$

or $N(a + b) \sin\theta_n = Nn\lambda,$ ……………………. (17)

Considering minima on either side of principal maxima,

then its direction is $(\theta_n \pm d\theta_n)$ with $d\theta_n \rightarrow$ angular half width of the nth maximum.

For the first minima of the nth principal maxima $m = Nn \pm 1$

$$N(a + b)\, sin(\theta_n \pm d\theta_n) = (Nn \pm 1)\lambda \quad \text{........(3)}$$

$$N(a + b)[sin\theta_n \cos d\theta_n \pm cos\theta_n\, sind\theta_n] = (Nn \pm 1)\lambda$$

Since $d\theta_n$ is very small, $\cos d\theta_n$ =1 and sindθ_n= $d\theta_n$

Thus $N(a + b)[sin\theta_n \pm cos\theta_n\, d\theta_n] = (Nn \pm 1)\lambda$

$$N(a + b)\, sin\theta_n \pm N(a + b)cos\theta_n\, d\theta_n = (Nn \pm 1)\lambda \quad \text{.....(4)}$$

Using condition (2) in (4) $\quad Nn\lambda \pm N(a + b)cos\theta_n\, d\theta_n = Nn\lambda \pm \lambda$

$$N(a + b)cos\theta_n\, d\theta_n = \lambda \quad or \quad d\theta_n = \frac{\lambda}{N(a + b)cos\theta_n}$$

The angular width of nth principal maxima is given by

$$\boldsymbol{(\theta_n + d\theta_n) - (\theta_n - d\theta_n) = 2d\theta_n = \frac{2\lambda}{N(a + b)cos\theta_n}}$$

<u>Maximum number of orders available with a grating</u>

For the principal maxima $(a + b)\, sin\theta = n\lambda$

or $\quad n = \dfrac{(a + b)\, sin\theta}{\lambda}$

The maximum angle of deflection is θ = 90^0, the maximum order is $n_{max} = \dfrac{(a + b)}{\lambda}$

If the grating element is less than twice the λ, then (a + b) $<\Box 2\lambda$ or $n_{max} < 2\,\lambda/\,\lambda <$ 2 . Thus only first order is possible.

10.6.2 Condition for Absent Spectra

Missing or absent spectra – In a plane diffraction grating, if the angle of diffraction θ is such that the minima due to diffraction component in the intensity distribution falls at the same positions of principal maxima due to interference component, then, that order of principal maxima will be missing or absent.

• It may be possible that while the first order spectra is clearly visible, second order may be not be visible at all and the third order may again be visible. It happen when for again angle of diffraction 0, the path difference between the diffracted ray from the two extreme ends of one slit is equal to an integral multiple of A if the path difference between the secondary waves from the corresponding point in the two halves will be A/2 and they will can all one another effect resulting is zero intensity. Thus the mining of single slit pattern are obtained in the direction given by.

$a \sin \theta = m\lambda$ .... (18)

where m = 1, 2, 3, ... excluding zero but the condition for nth order principles maximum in the grating spectrum is

$(a + b) \sin \theta = n\lambda$ .... (19)

If the two conditions given by equation (19) are simultaneously satisfied then the direction in which the grating spectrum should give us a maximum every slit by itself will produce darkness in the direction and hence the most favourable phase for reinforcement will not be able to produce an illumination i.e., the resultant intensity will be zero and hence the absent spectrum. Therefore dividing equation (19) by equation (18).

$((a+ b) \sin \theta)/a \sin \theta = n/ m$

$(a+ b) /a = n/m$

$(a + b) \sin \theta = n\lambda$ ....(2)

- If the two conditions given by equation (2) are simultaneously satisfied then the direction in which the grating spectrum should give us a maximum every slit by itself will produce darkness in that direction and hence the most favourable phase for reinforcement will not be able to produce an illumination i.e., the resultant intensity will be zero and hence the absent spectrum. Therefore dividing equation (2) by equation (1)

$((a+ b) \sin \theta)/a \sin \theta = n/ m$

$(a+ b) /a = n/m$

- This is the condition for the absent spectra in the diffraction pattern If a= b i.e., the width of transparent portion is equal to the width of opaque portion then from equation (3) n = 2m i.e., 2nd, 4th, 6th etc., orders of the spectra will be absent corresponds to the minima due to single slit given by m = 1, 2, 3 etc.

b = 2a

n=3m

i.e., 3rd, 6th, 9th etc., order of the spectra will be absent corresponding to a minima due to a single slit given by m = 1, 2, 3 etc.

In a plane diffraction grating, if the angle of diffraction is such that the minima due to diffraction component in the intensity distribution falls at the same positions of principal maxima due to interference component, then, that order of principal maxima will be missing or absent.

10.7 WHAT IS GRATING SPECTRA

Grating spectra is the spectra we can obtain from grating prisms. These spectra appear as line spectra, and they form due to the diffraction of light. This technique is very important in analyzing light sources. A grating spectrum contains a large number of equally spaced parallel slits. The basic phenomenon of the working principle of the grating spectra is light diffraction. There are spaces between lines of this spectrum that appears as slits; these slits diffract the light waves, producing many different beams that can interfere to produce a spectrum.

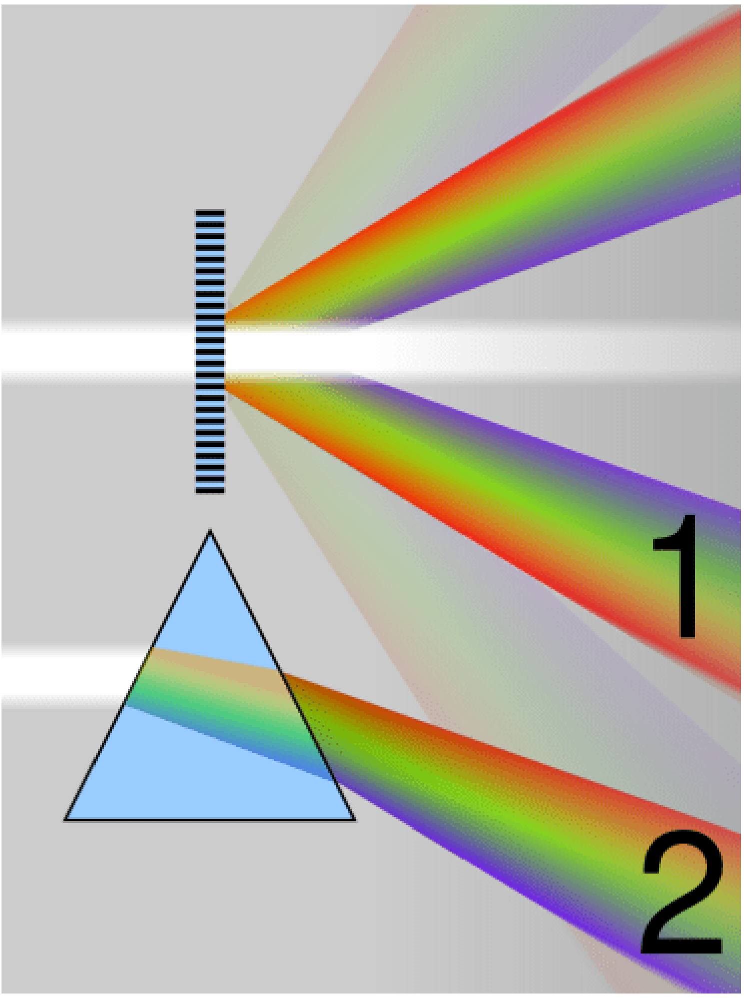


Figure : 9

A grating prism or grism can be explained as a combination of a prism and a grating system that is arranged along with prims, which allows the light of a chosen wavelength to pass through the prism straightly. This prism system has the advantage of allowing a single camera to be used for imaging and spectroscopic needs without removing or changing the prism.

10.8 WHAT IS PRISM SPECTRA

Prism spectra are the continuous spectra we can get using a prism. A prism is a transparent instrument that is triangular and has a refracting medium that can cause the refraction of light. It has a base and an apex, and its apical angle tends to determine the diprotic power of the material. When light passes through a prism, the light gets dispersed by it, giving a prism spectrum.

Visible light is usually white light, which contains a collection of component colours. Often, we can observe these colours when the light passes through a triangular prism. This is because the white light separates into its component colours when the light passes through the prism. The colour components that we can observe are red, orange, yellow, green, blue, and violet. Typically, this colour separation process is known as dispersion.

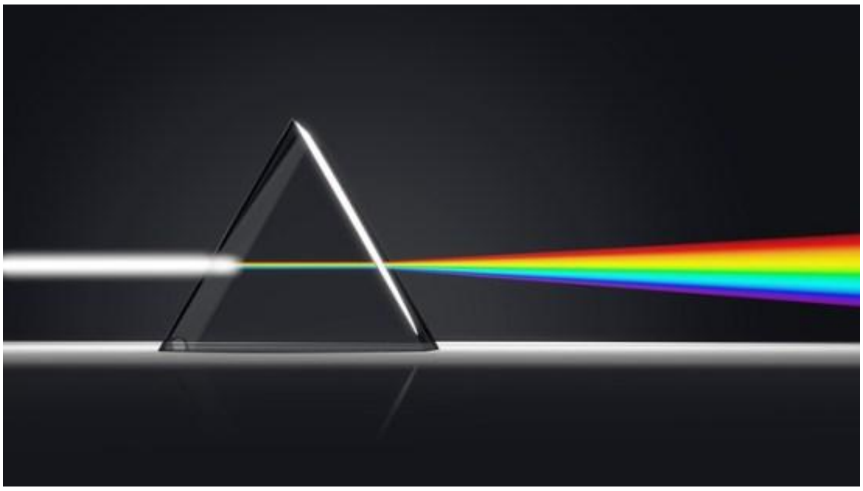


Figure : 10

The dispersion of colours in light can be described based on the varying frequencies and wavelengths of each colour component. These different frequencies of light tend to bend in varying amounts as the light passes through the prism.

When considering the material of the prism, different materials have different optical densities (the optical density is the tendency of a material to slow down the light when the light passes through that material). When light is passing through a transparent material, it tends to interact with the atoms of the material. If the frequency of the light wave matches the resonance frequency of the electrons in atoms, the light is absorbed by that atom. The unabsorbed light comes out of the prism, which gives us the prism spectrum.

10.9 THEORY OF CONCAVE DIFFRACTION GRATING

Let GG' be a concave grating having of curvature at C (in below fig.). Let A and B be two corresponding points of the grating so that AB = (e+ d) is the grating element. Let S be a narrow slit perpendicular to the plane of the paper and illuminated by monochromatic light of wavelength. Let SA and SB be two rays incident on the grating at angle i and i + di respectively. AS 'and BS' are the corresponding diffracted rays with angel of diffraction and respectively

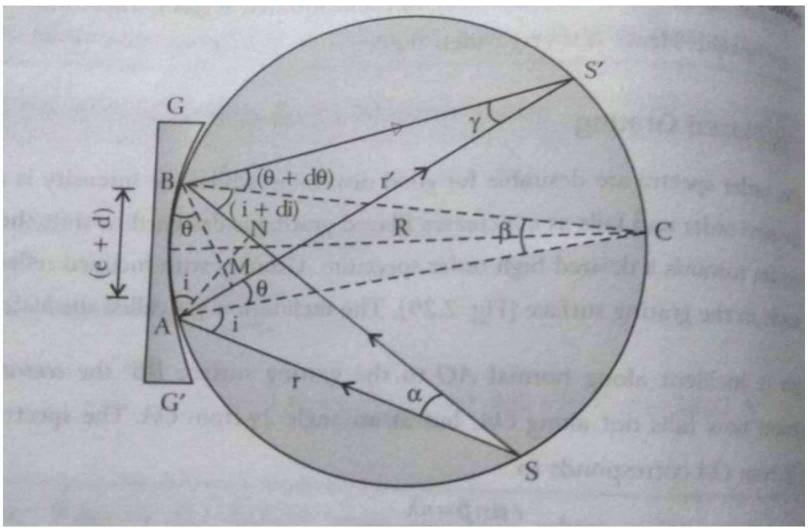


Figure : 11

=(SB+BS')-(SA+AS')

=(SB-SA)-(AS'-BS')

=BN-AM

=AB(sin i-AB sin θ)

= (e+d)(sini-sin θ)

[AB =(e+d)]

For nth order spectrum, a maximum at S' occurs only when

(e+d)(sini- sin θ) =n

In order to from a diffraction image at S', the path defference between any two consecutive rays should be constant i.e.,

=(e+d)(sin i- sin θ) =constant

On differentiating we get,

(cos i di- sin θ d) =0

Pitting these values in we get,

But, from above fig. we have

$\alpha + i = \beta + i + di$ $\quad \therefore di = \alpha - \beta$

$\beta + \theta = \gamma + \theta + d\theta$ $\quad \therefore d\theta = \beta - \gamma$

$$\cos i(\alpha - \beta) - \cos\theta(\beta - \gamma) = 0$$

Let SA=r, AS'=r' and radius of curvature of the grating be R, then

$$r.\alpha = AN = (e+d)\ \cos i$$

$$R\beta = AB = (e+d)$$

$$r'\gamma = BM = (e+d)\cos\theta$$

$$\therefore \quad \alpha = (e+d)\cos i, \quad \beta = \frac{e+d}{R}, \quad and\ \gamma = \frac{(e+d)\cos\theta}{r'}$$

Putting these values in we get,

$$\cos i \left[\frac{\cos i}{r} - \frac{1}{R}\right] - \cos\theta\left[\frac{1}{R} - \frac{\cos\theta}{r'}\right] = 0$$

It is the general eq. for the position of S' which shows that,

If, r=R cos i than r'=R $\cos\theta$

i.e., If S lies at the circumference of a circle of radius R, then S' also lies on the same circle. Thus, we can say that if the slit and the concave grating are placed at the 9 circumference of a circle whose diameter is equal to the radius of curvatures of the grating, then the spectra are focused on the circumference of the same circle.

10.9.1 Type of Concave Diffracting Grating

In general, there are four types of diffraction gratings: **ruled gratings, holographic gratings, transmission gratings, and reflection gratings**. Ruled gratings are created by physically etching several parallel grooves onto a reflective surface.

10.10 SUPERIORITY OVER PLANE GRATING

Spectroscopy, like most sciences, never stops evolving. Revolutionary techniques, solutions, and instruments are devised every year. With improvements to grating design and optical components, spectroscopy has become an even more powerful tool used to analyze the world around us.

The use of plane gratings in spectroscopy is still popular due to their efficiency and ease of production. However, instruments utilizing plane gratings also require the use of mirrors and other optics, which can introduce aberrations and produce weaker sensitivity at ultraviolet and near-infrared wavelengths.

Most plane grating optics utilize several mirrors in order to properly image the spectrum on the detector. With applications that require the use of UV light, especially UV-Vis absorbance, fluorescence, and solar radiometry, multiple mirrors can be detrimental to the performance of an instrument. In some plane grating designs, UV sensitivity can be weak. The use of a holographic concave grating can eliminate the sensitivity issues and provide other benefits including: stray light reduction, aberration correction, thermal stability, and improved ruggedness.

The advantage of a concave grating over a plane grating is its ability to produce sharp spectral lines without the aid of lenses or additional mirrors. This makes it useful in the infrared and ultraviolet regions in which these radiations would otherwise be absorbed upon passage through a lens.

Figure : 12

10.11 RAYLEIGH'S CRITERION

Two bright point objects lying very close to each other are seen with a naked eye or with the help of an optical instrument (a microscope or a telescope), then they may or may not be seen as two separate distinct objects due to the overlapping of their diffraction patterns. The ability to separate distinct objects depends upon limit of resolution and the resolving power of an optical instrument.

The smallest linear or angular separation between two point objects at which they can be just separately seen or resolved by an optical instrument is called the limit of resolution of the instrument.

The resolving power of an optical instrument is its ability to resolve or separate the images of two nearby point objects so that they can be distinctly seen. It is equal to the reciprocal of the limit of resolution of the optical instrument.

The smaller the limit of resolution of an optical instruments, greater is its resolving power.

10.11.1 Rayleigh Criterion of Resolution

Rayleigh's criterion for resolution : If we look through a telescope at two stars lying close to each other, their diffraction patterns overlap and the resultant pattern is little broader but otherwise similar to that of a single star, as shown in figure. So the two stars cannot be resolved or separately seen.

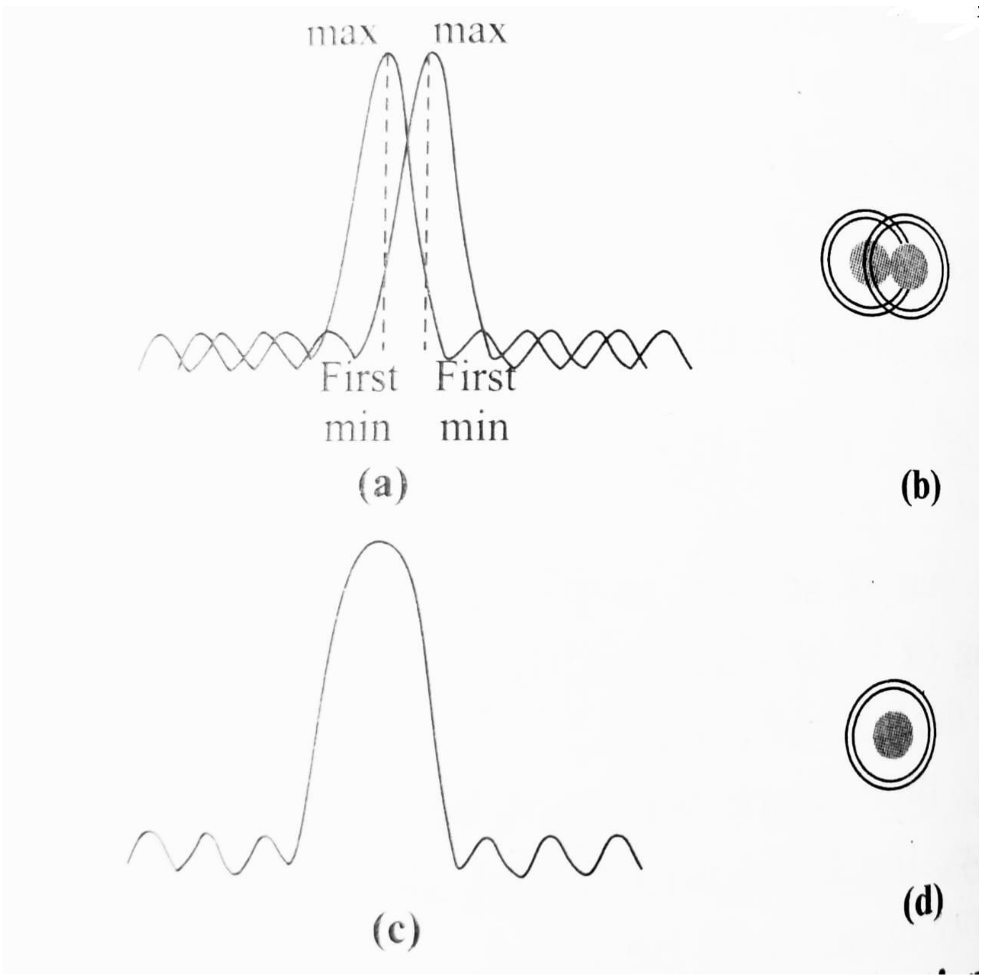


Figure : 13

Figure 13 (a) and (b) show overlapping of diffraction patterns of two close point objects and Fig. (c) and (d) show their resultant diffraction patterns.

According to Rayleigh criterion, the images of two point objects are just resolved when the central maximum of the diffraction pattern of one falls over the first minimum of the diffraction pattern of the other, as shown in figure. When seen through the telescope, the resultant diffraction has a well-marked depression at the top, showing that these are really two stars and not one. Thus the images of two stars have been just resolved.

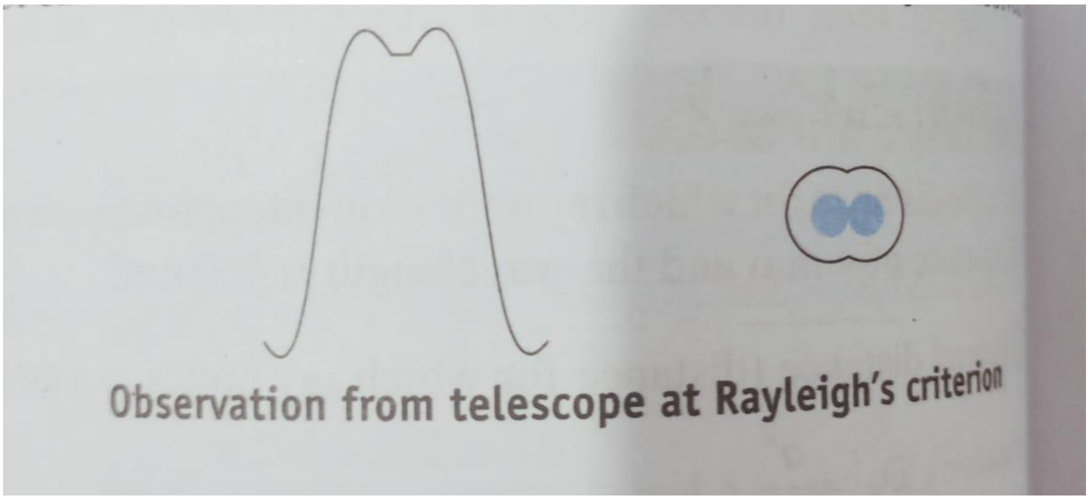


Figure : 14

10.11.2 Limits of Resolution of Eye

The resolution of the eye is the smallest object that can be seen by the human eye. The eye has a small aperture called a pupil which controls the amount of light that enters our eyes. The resolution of the eye is dependent on the pupil size which would imply that our eyes could resolve more detail in the dark when the pupil is large than in bright light when it is small.

The human eye's natural pupil size is 2 mm, which determines the minimum angular resolution of the eye and enables us to see tiny objects as close as possible to our eyes, although there is a minimum distance of about 25 cm for comfortable vision. But the number quoted is 0.1 mm for the smallest resolvable dimension, indicating that the diffraction limit is a key factor in the strength of visual resolution.

We can also say that the limit of resolution of the eye is roughly 1 minute (angle).

10.12 RESOLVING POWER

The ability of an optical instrument to show two close lying point objects as well separated point objects is called its **resolving power**. The resolution is limited by the diffraction patterns of the two close lying point objects which overlap as shown.

10.12.1 Resolving Power of Grating

The resolving power of a grating is **a measure of its ability to spatially separate two wavelengths**. It is determined by applying the Rayleigh criteria to the diffraction maxima; two wavelengths are resolvable when the maxima of one wavelength coincides with the minima of the second wavelength.

It is the capacity of the grating to form separate diffraction maxima of two wavelengths that are close to each other. The direction of nth principal maximum for wavelength λ is

The equation for minima is $N(a + b)\, sin\theta_n = m\lambda$ where m has all integral values except $0,\ N,\ 2N.....nN$ because for these values of m, the condition for maxima is satisfied.

Thus the first minimum adjacent to n[th] principal maximum in the direction $\theta_n + d\theta$ is obtained by substituting the value of m as $(nN + 1)$. Thus the first minimum in the direction of $(\theta_n + d\theta)$ is $N(a + b)\sin(\theta_n + d\theta) = (nN + 1)\lambda$ .....(2)

The direction of n[th] principal maximum for wavelength λ+d λ is

$(a + b)\, sin(\theta_n + d\theta) = n(\lambda + d\lambda)$ ....(3)

Multiplying (3) by N, we have $N(a + b)\, sin(\theta_n + d\theta) = nN(\lambda + d\lambda)$ ....(4)

The two lines appear just resolved if the angle diffraction $(\theta_n + d\theta)$ also correspond to the direction of first secondary minimum due to the first diffraction pattern.

Comparing (2) and (4)

$nN(\lambda + d\lambda) = (nN + 1)\lambda$ or $n(\lambda + d\lambda) = n\lambda + \frac{\lambda}{N}$

$n\, d\lambda = \frac{\lambda}{N}$ or $\frac{\boldsymbol{\lambda}}{\boldsymbol{d\lambda}} = \boldsymbol{nN}$ → Expression for resolving power of grating

Thus the resolving power is (1) directly proportional to the order of the spectrum and (2) the total number of lines on the grating surface.

$$(a + b)\, sin\theta_n = n\lambda \quad (1)$$

10.12.2 Prism

Ability of a prism to separated two close spectral lines is known as resolving power of the prism. It is measured by $\frac{\lambda}{d\lambda}$, where d $\boldsymbol{\lambda}$ is the smallest wavelength difference that can be just resolved by the wavelength λ**.**

Let a plane wave front BP of light wavelength λ and λ +d λ be incident on the prism ABC which is placed on position of minimum deviation. Let CQ and CQ' be the emergent wavefront for wavelength λ and λ +d λ as shown in below fig. let δ and $\delta + d\delta$ be the angle of deviation, μ and $\mu + d\mu$ these refractive induces corresponding to wavelength λ and λ +d λ respectively.

According to Fermat's principle, the optical path between the incident and emergent wavelength for any wavelength must be the same. Hence for the wavelength λ.

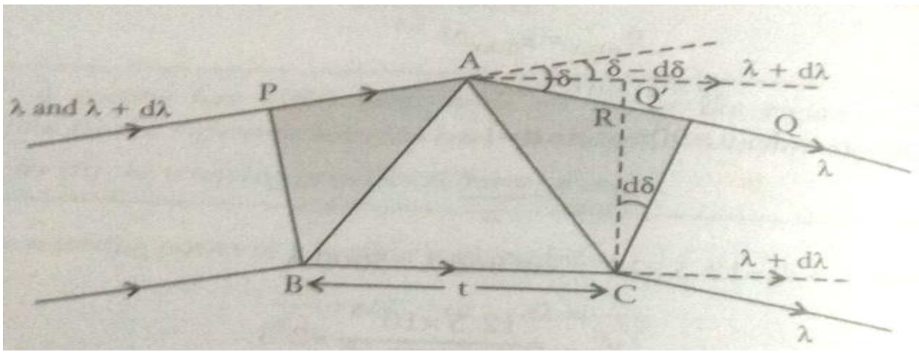


Figure : 15

$$PA+AQ= \mu\ BC \qquad \ldots\ldots (1)$$

Similarly, for the wavelength $\lambda + d\lambda$

$$PA+AQ'= (\mu - d\mu)BC \qquad \ldots. (2)$$

From eq.(1) and (2), we have

$$AQ-AQ'= d\mu\ BC$$

but from above fig.

AQ=AR, so that AQ-AQ'=AQ-AR=RQ

$$RQ= d\mu BC$$

Again, from fig.

$$Q'CQ = d\delta$$

$$RQ=CQ\ d\delta$$

$$CQ\ d\delta = d\mu\ BC$$

The emergent beam has rectangular section. The section may also be considered as a rectangular aperture of width e which is the width of the emergent be CQ.

Hence,

$$e d\delta = t\ d\mu$$

$$d\delta = \frac{t d\mu}{e} \qquad (3)$$

Where, t is the width of the base of prism.

Hence, the diffraction pattern for each wavelength is equivalent to that due to a rectangular aperture of width e. according to Rayleigh criterion, if two spectral lines are just resolved, then principal maximum of one should fall on the principle maximum of the other. it means that angular separation $d\delta$ betaween the principal maxima of λ

and λ +d λ must be equal to the angle $d\theta$ between the principal maximum and first minimum of λ. The angle $d\theta$ *is given* by the Fraunhofer diffraction at a single slit.

$$e \sin d\theta = \lambda$$

$$e d\theta = \lambda$$

$$d\theta = \frac{\lambda}{e} \qquad \ldots\ldots\ldots (4)$$

For just resolution $d\delta = d\theta$. So, from eq.(3) and (4), we get

$$\frac{t d\mu}{e} = \frac{\lambda}{e}$$

$$t d\mu = \lambda$$

Hence, the resolving power of the prism,

$$\frac{\lambda}{d\lambda} = t\frac{d\mu}{d\lambda}$$

It show that the resolving power of a prism is directly proportional to the width of the base of the prism and also proportional to the rate of change of refractive index with wavelength. As we have $\frac{d\mu}{d\lambda}$= **dispersive power.**

So, **Resolving power= t* dispersive power**

10.12.3 Telescope

In telescopes, very close objects such as binary stars or individual stars of galaxies subtend very small angles on the telescope. To resolve them we need very large apertures.

In above fig. the image of each point object is a Fraunhofer diffraction Patten and lies in the focal plane of the telescope objective AB. Let P1 and P2 be the position of centre of maxima of the two images. We can use Rayleigh's to determine the resolving power. The angular separation between two objects must be

The path difference will be equal to,

$$BP2\text{-}AP2=\lambda$$

But from fig.

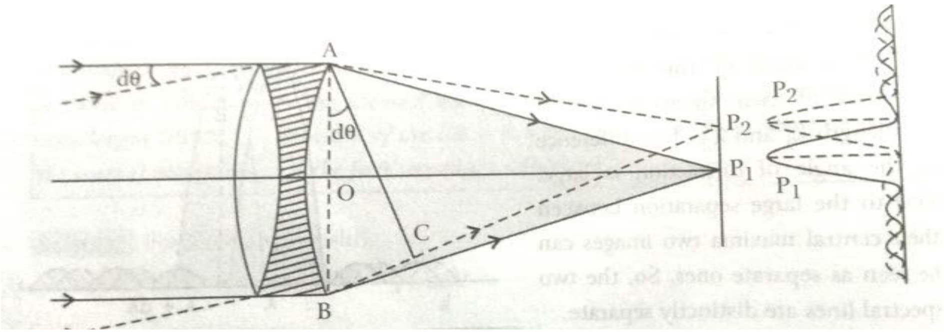


Figure : 16

$$BP2\text{-}AP2=BC=AB\ d\theta$$

[the angle is small, sinθ= dθ]

$=D\ d\theta$

[AB=D=diameter of objective lens of telescope]

$$d\theta=\lambda/D$$

for circular apertures, Airy showed that 1.22 λ may be used instead of λ .

$$\mathbf{d\theta = 1.22\,\lambda\ /D}$$

Thus higher the diameter *D*, better the resolution. The best astronomical optical telescopes have mirror diameters as large as 10m to achieve the best resolution. Also, larger wavelengths reduce the resolving power and consequently radio and microwave telescopes need larger mirrors.

Relation between magnifying power and the resolving power of telescope:

From the theory of telescope, we know that magnifying power of a telescope is given by

M= Diameter of objective (entrance pupil)/diameter of exit pupil (eye ring)

= D/d ……………..(1)

The magnifying power is said to be normal if diameter d of the eye ring is equal to the diameter d_e, of the pupil of eye. Hence,

$$\text{Normal magnifying power} = \frac{\text{Diameter of the objective}}{\text{Diameter of the eye}} = \frac{D}{d_e}$$

Further, the limit of resolution of the telescope objective of diameter D is given by

$$d\theta = \frac{1.22\lambda}{D} \qquad \dots(2)$$

and the limit of resolution of the unaided eye is given by

$$d\theta' = \frac{1.22\lambda}{d_e} \qquad \dots(3)$$

$$\because \frac{\text{Limit of resolution of the eye}}{\text{Limit of resolution of the telescope}} = \frac{d\theta'}{d\theta} = \frac{1.22\lambda}{d_e} \Big/ \frac{1.22\lambda}{D}$$

$$= \frac{D}{d_e} = \text{Normal magnifying power of the telescope} \quad \dots(4)$$

Thus, <u>the limit of resolution of telescope, multiplied by its normal magnifying power is equal to the limit of resolution of the unaided eye.</u>

If the wavelength of light is taken as 5000 Å and the pupil diameters of the eye as 2 mm, the minimum resolving angle between two distant point objects by the eye

$$d\theta' = \frac{1.22\lambda}{d_e} = \frac{1.22 \times 5000 \times 10^{-8}}{0.2} \text{ radians}$$

$$= \frac{1.22 \times 5 \times 10^{-5}}{0.2} \times \frac{180}{\pi} \times 60 \text{ minutes}$$

=1 minute of arc = 60 seconds of arc.

Similarly the resolvable angle of the 40 inch refracting telescope

$$d\theta' = \frac{1.22\lambda}{D} = \frac{1.22 \times 5 \times 10^{-5}}{40 \times 2.54} \times \frac{180}{\pi} \times 60 \times 60 \text{ sec.}$$

$$= \frac{5}{40} = \frac{1}{8} \text{ second}$$

Thus, we can say that the least angular separation (in seconds) resolvable by a telescope is

$$\frac{5}{\text{aperture in inches}}$$

which is Dawe's rule.

Hence, the normal magnifying power of the telescope with 40 inch objective

$$= \frac{d\theta'}{d\theta} = \frac{60}{1/8} = 480.$$

Thus to bring out the full details of the object, the magnifying power of about 480 is required. If the magnifying power is greater than this, it will serve no useful purpose because it would not be accompanied by any more resolution and if it is less than this, the full advantage of high resolving power of the telescope will not be taken.

10.12.4 Microscope

For microscopes, the resolving power is the inverse of the distance between two objects that can be just resolved.

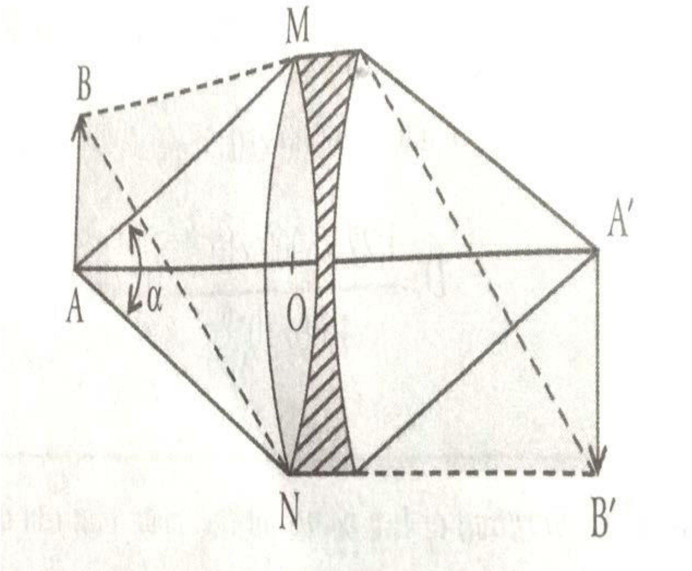


Figure : 17

In fig.(17), MN is the aperture of the objective of a microscope and A and B are two object points at a distance d apart. A' and B' are the corresponding Fraunhofer diffraction pattern of the two images. A' is the position of the central maximum of A and B' is the position of the central maximum of B'. the two image are said to be just resolved if the position of the central maximum B' also correspond to the first minimum of the image of A'.

$$(BN-NA') – (BM-MA')$$

$$NA'=MA'$$

Path difference = BN-BM

In fig.( ), AD is perpendicular to DM and AC is perpendicular to BN.

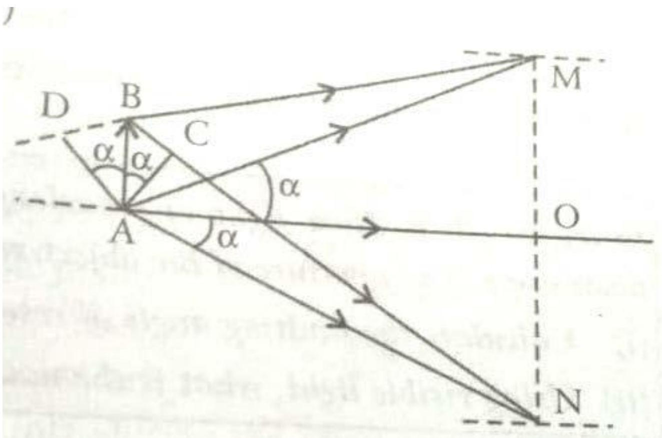


Figure : 18

$$BN-BM = (BC+CN)-(DM-DB)$$

$$CN=AN=AM=DM$$

Path difference=BC+DB

From the Δ ACB and Δ ADB

$$BC=AB \sin\alpha = d \sin\alpha$$

$$DB=AB \sin\alpha = d \sin\alpha$$

Path difference=2d $\sin\alpha$

If this Path difference 2d $\sin\alpha$=1.22 $\boldsymbol{\lambda}$, then, A' corresponds to the first minimum of the image B' and the two image appear just resolved.

$$2d \sin\alpha = 1.22\, \boldsymbol{\lambda}$$

$$\mathbf{d=1.22\, \boldsymbol{\lambda} /2 \sin\alpha}$$

Where *n* is the refractive index of the medium separating object and aperture. Note that to achieve high-resolution *n sin θ* must be large. This is known as the Numerical aperture.

Thus, for good resolution:

1. sinθ must be large. To achieve this, the objective lens is kept as close to the specimen as possible.
2. A higher refractive index (n) medium must be used. Oil immersion microscopes use oil to increase the refractive index. Typically for use in biology studies, this is

limited to 1.6 to match the refractive index of glass slides used. (This limits reflection from slides). Thus the numerical aperture is limited to just 1.4-1.6. Thus, optical microscopes (if you do the math) can only image to about 0.1 microns. This means that usually organelles, viruses and proteins cannot be imaged.

3. Decreasing the wavelength by using X-rays and gamma rays. While these techniques are used to study inorganic crystals, biological samples are usually damaged by x-rays and hence are not used.

10.13 SUMMARY

The basics of the diffraction phenomena along with various classes of diffraction have been discussed. The Fraunhofer diffraction for single slit, double slit, circular aperture and N slits (grating) have been discussed in the details. The calculation for the intensity of the principal maxima, secondary maxima and minima has been derived. Their relative comparison in terms of their intensities has also been made. Determination of missing orders in case of double slit and N slits (grating) diffraction pattern has also been made.

The resolving power of an optical instrument is defined as its ability to just resolve the images of two close point sources or small objects. The Rayleigh Criterion gives a quantitative account of the phenomena of resolution. The definitions and physical meanings for the resolving powers of diffraction grating, prism, telescope and microscope were discussed. Their mathematical expressions have also been derived in the present chapter.

10.14 TERMINAL QUESTIONS

1. Describe and Explain Fraunhofer diffraction.
2. Discuss Fraunhofer diffraction due to a Single Slit.
3. Discuss the diffraction of Light by a narrow slit.
4. What is plane diffraction grating?
5. Distinguish between grating spectra and Prism Spectra.
6. Discuss Rayleigh criterion for resolution. What is limit of resolution?
7. Define the resolving power of microscope.

10.15 ANSWER AND SOLUTIONS OF TERMINAL QUESTION

1. Section 10.3
2. Section 10.4

3. Section 10.5

4. Section 10.6

5. Section 10.7 & 10.8

6. Section 10.11 & 10.11.1

7. Section 10.12.3

10.16 SUGGESTED READINGS

1. Elementary Wave Optics, H. Webb Robert, Dover Publication.

2. Optics – Ajoy Ghatak

3. Optics – Principles and Application, K. K. Sharma, Academic Press.

4. Introduction to Optics – Frank S. J. Pedrotti, Prentice Hall.

5. Advanced Geometrical Optics – Psang Dain Lin.