Chapter Four

Interference

CHAPTER OUTLINE

4.1 Superposition of Waves

4.2 Interference

4.2.1 Theory of Interference

4.2.2 Intensity Distribution

4.2.3 Superposition of Incoherent Waves

4.2.4 Superposition of Many coherent Waves

Michael Faraday

British Physicist and Chemist (1791–1867)

Faraday is often regarded as the greatest experimental scientist of the 1800s. His many contributions to the study of electricity include the invention of the electric motor, electric generator, and transformer, as well as the discovery of electromagnetic induction and the laws of electrolysis. Greatly influenced by religion, he refused to work on the development of poison gas for the British military. *(By kind permission of the President and Council of the Royal Society)*

4.1 Superposition of Waves

Many interesting wave phenomena in nature cannot be described by a single traveling wave. Instead, one must analyze complex waves in terms of a combination of traveling waves. To analyze such wave combinations, one can make use of the **superposition principle:**

If two or more traveling waves are moving through a medium, the resultant value of the wave function at any point is the algebraic sum of the values of the wave functions of the individual waves.

The principle of superposition applies to electromagnetic wave also and is the most important principle in wave optics. In case of electromagnetic waves, the term displacement refers to the amplitude of electric field vector.

Interference is an important consequence of superposition of coherent waves.

Waves that obey this principle are called ***linear waves***. Waves that violate the superposition principle are called ***nonlinear waves***.

4.2 Interference

If two or more light waves of the same frequency overlap at a point, the resultant effect depends on the phase of the waves as well as their amplitudes. The resultant wave at any point at any instant of time is governed by the principle of superposition. The combined effect at each point of the region of superposition is obtained by adding algebraically the amplitudes of the individual waves. Let us assume here that the component waves are of the same amplitude.

At certain points, the two waves may be in phase. The amplitude of the resultant wave will then be equal to the sum of the amplitudes of the two waves, as shown figure 1

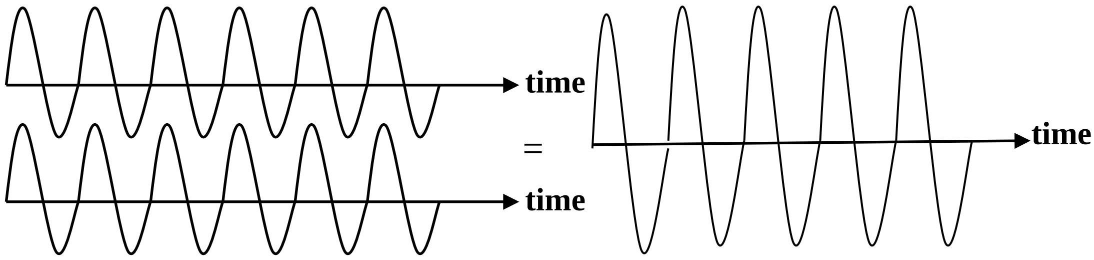

Figure 1

Thus, the amplitude of resultant wave

A_R = A+A = 2A. (4.1)

Hence, the intensity of the resultant wave

$I_R \propto A_R^2 = 2A^2 = 2^2 I.$ (4.2)

It is obvious that the resultant intensity is greater than the sum of intensities due to individual waves.

$I_R > I + I = 2I$ (4.3)

Therefore, the interference produced at these points is known as ***constructive interference.*** A stationary bright band of light is observed at points constructive interference.

At certain other points, the two waves may be in opposite phase. The amplitude of the resultant wave will then be equal to the sum of the amplitudes of the two waves, as shown figure 2

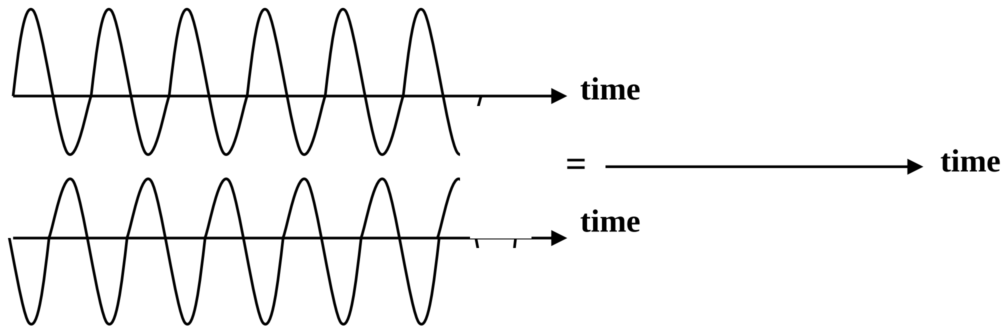

Figure 2

Thus, the amplitude of resultant wave

A_R = A-A = 0. (4.4)

Hence, the intensity of the resultant wave

$I_R \propto 0^2 = 0.$ (4.5)

It is obvious that the resultant intensity is greater than the sum of intensities due to individual waves.

$I_R < 2I$ (4.6)

Therefore, the interference produced at these points is known as ***destructive interference.*** A stationary dark band of light is observed at points destructive interference.

$\Delta = m\lambda$ $\boldsymbol{constructive\ interference}$ (4.7)

$\Delta = (2m+1)\frac{\lambda}{2}$ $\boldsymbol{destructive\ interference}$ (4.8)

Where m=0,1,2,3,……………….. and Δ optical path difference.

the phenomenon of redistribution of light energy due to the superposition of light waves from two or more coherent sources is known as interference.

4.2.1 Theory of Interference

Analytical Method: let us assume that the electric field component of the two waves arriving at point P vary with time as

$$E_A = E_1 \sin(w\,t) \tag{4.9}$$

$$E_B = E_2 \sin(w\,t + \delta) \tag{4.10}$$

According to principle of superposition obtain to

$$E_R = E_A + E_B = E_1 \sin(w\,t) + E_2 \sin(w\,t + \delta)$$

$$E_R = E_1 \sin(w\,t) + E_2\,(\sin w\,t \cos\delta + \cos w\,t \sin\delta)$$

$$E_R = (\,E_1 + E_2 \cos\delta\,) \sin(w\,t) + E_2\,\sin\delta \cos w\,t \tag{4.11}$$

Equation (4.11) shows that the ***superposition of two sinusoidal the same frequency but with a phase difference produces a sinusoidal wave with the same frequency but with a different amplitude*** **E.**

Let $$E_1 + E_2 \cos\delta = E \cos\Phi \tag{4.12}$$

and $$E_2\,\sin\delta = E \sin\Phi \tag{4.13}$$

where E is the amplitude of the resultant wave and Φ is the new initial phase angle. In order to solve for E and Φ, we square equation (4.12) and (4.13) and add them.

$$(E_1 + E_2 \cos\delta\,)^2 + E_2^2 \sin^2\delta = E^2(\cos^2\Phi + \sin^2\Phi)$$

$$E^2 = E_1^2 + E_2^2 \cos^2\delta + 2E_1E_2 \cos\delta \quad + E_2^2 \sin^2\delta$$

$$E^2 = E_1^2 + E_2^2(\cos^2\delta + \sin^2\delta) + 2E_1E_2 \cos\delta$$

$$E^2 = E_1^2 + E_2^2 + 2E_1E_2 \cos\delta \tag{4.14}$$

Thus, it is seen that the square of the amplitude of the resultant wave is not a simple sum of the squares of the amplitudes of the superposition waves, there is an additional term which is known as the interference term.

4.2.2 Intensity Distribution

The intensity of a light wave is given by the square of its amplitude.

$$I = \frac{1}{2}\epsilon_{\circ} C E^2 \propto E^2$$

Using this equation into (4.14), we get

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta \qquad (4.15)$$

The term $2\sqrt{I_1 I_2}\cos\delta$ is known as the ***interference tern***. Whenever the phase difference between the waves is zero, i.e. $\delta = 0$, we have maximum amount of light. Thus,

$$I_{max.} = I_1 + I_2 + 2\sqrt{I_1 I_2} \qquad (4.16)$$

When $I_1 = I_2 = I_{\circ}$

$$I_{max.} = 4I_{\circ} \qquad (4.16a)$$

When the phase difference is $\delta = 180^{\circ}, \cos 180^{\circ} = -1$ and we have minimum amount of light.

$$I_{min.} = I_1 + I_2 - 2\sqrt{I_1 I_2} \qquad (4.17)$$

Which, when $I_1 = I_2$ becomes

$$I_{min.} = 0 \qquad (4.17a)$$

At point that lie between the maxima and minima, when $I_1 = I_2 = I_{\circ}$, we get

$$I = I_{\circ} + I_{\circ} + 2I_{\circ}\cos\delta = 2I_{\circ}(1 + \cos\delta)$$

The using the identity $1 + \cos\delta = 2\cos^2\left(\frac{1}{2}\delta\right)$, we get

$$I = 4I_{\circ}\cos^2\left(\frac{1}{2}\delta\right) \tag{4.18}$$

This equation shows that the intensity varies along the screen in accordance with the ***law of cosine square***.

4.2.3 Superposition of Incoherent Waves

Incoherent waves are the waves that do not maintain a constant phase difference. Then the phase fluctuate irregularly with time and independently of each other. In case of light waves the phase fluctuate randomly at a rate of about 10^8 per second. Light detectors such as human eye, photographic film etc cannot respond to such rapid change. The detected intensity is always the average intensity, averaged over a time interval which is very much larger than the time of fluctuation. Thus,

$$I_{ave} = I_1 + I_2 + 2\sqrt{I_1 I_2}\langle\cos\delta\rangle$$

The average value of the cosine over a large time interval will be zero and hence the interference term become zero. Therefore, the average intensity of the resultant wave is

$$I_{ave} = I_1 + I_2$$

If $I_1 = I_2$, then

$$I_{ave} = 2I \tag{4.19}$$

It implies that the superposition of incoherent waves does not produce interference but gives a uniform illumination. The average intensity at any point is simply equal to the sum of the intensities of the component waves.

4.2.4 Superposition of Many coherent Waves

The result (4.16) may be written as

$$I_{max.} = 2^2 I_\circ$$

Which gives the resultant intensity when two coherent waves superpose. The resultant maximum intensity due to N coherent waves will be therefore

$$I_{max.} = N^2 I_\circ \qquad (4.20a)$$

$$I_{min.} = 0 \qquad (4.20b)$$

Where N represent the number of coherent waves superposition at a point.

4.3 Huygens's Principle

Huygens's principle is a geometric construction for using knowledge of an earlier wave front to determine the position of a new wave front at some instant. In Huygens's construction,

the principle states that every point of a wave front may be considered as the source of small secondary wavelets.

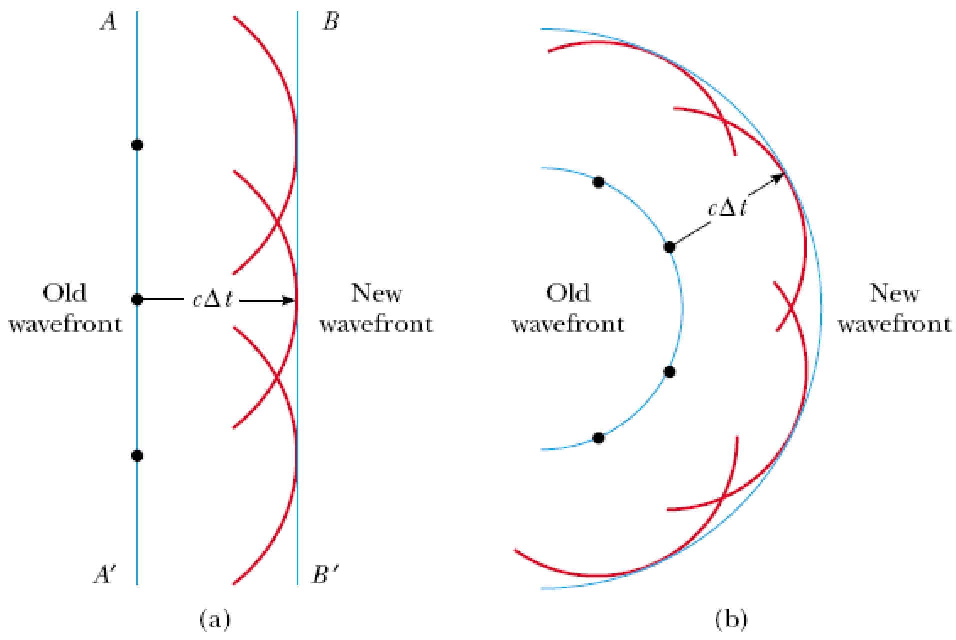

Figure 3: Huygens's construction for (a) a plane wave propagating to the right and (b) a spherical wave propagating to the right.

4.4 Young's Double-Slit Experiment-Wave front Division

A common method for producing two coherent light sources is to use a monochromatic source to illuminate a barrier containing two small openings (usually in the shape of slits). The light emerging from the two slits is coherent because a single source produces the original light beam and the two slits serve only to separate the original beam into two parts . Any random change in the light emitted by the source occurs in both beams at the same time, and as a result interference effects can be observed when the light from the two slits arrives at a viewing screen. If the light traveled only in its original direction after passing through the slits, as shown in Figure 4a, the waves would not overlap and no interference pattern would be seen. the waves spread out from the slits as shown in Figure 4b. In other words, the light deviates from a straight-line path and enters the region that would otherwise be shadowed.

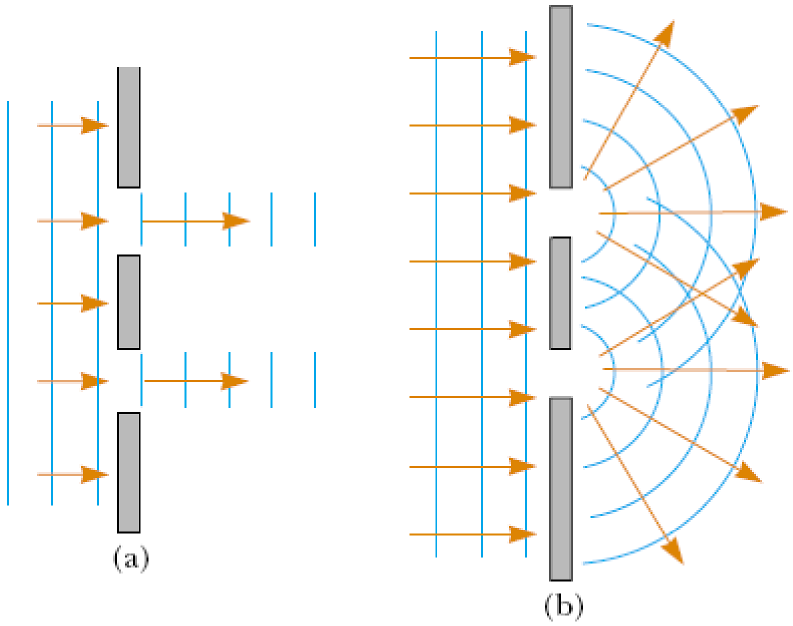

Figure 4 (a) If light waves did not spread out after passing through the slits, no interference would occur. (b) The light waves from the two slits overlap as they spread out, filling what we expect to be shadowed regions with light and producing interference fringes on a screen placed to the right of the slits.

Interference in light waves from two sources was first demonstrated by Thomas Young in 1801. A schematic diagram of the apparatus that Young used is shown in Figure 5a. Plane light waves arrive at a barrier that contains two parallel slits S_1 and S_2. These two slits serve as a pair of coherent light sources because waves emerging from them originate from the same wave front and therefore maintain a constant phase relationship. The light from S_1 and S_2 produces on a viewing screen a visible pattern of bright and dark parallel bands called fringes (Figure 5b). When the light from S_1 and that from S_2 both arrive at a point on the screen such that constructive interference occurs at that location, a bright fringe appears. When the light from the two slits combines destructively at any location on the screen, a dark fringe results.

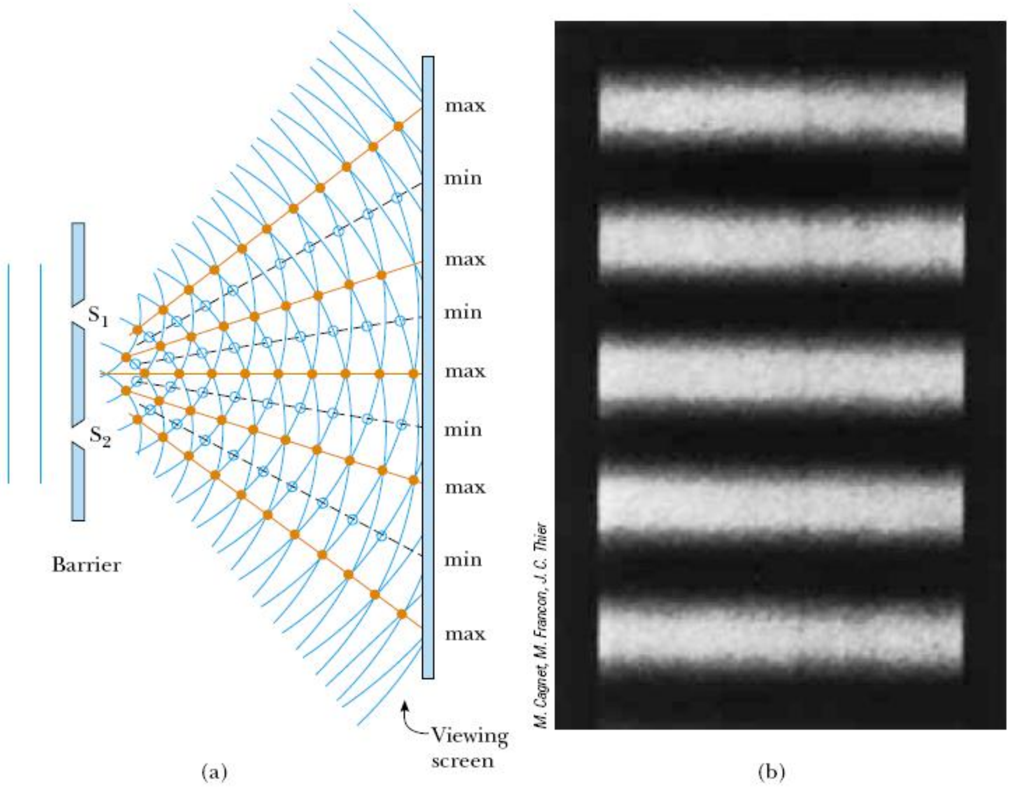

Figure 5 (a) Schematic diagram of Young's double-slit experiment. Slits S1 and S2 behave as coherent sources of light waves that produce an interference pattern on the viewing screen (drawing not to scale). (b) An enlargement of the center of a fringe pattern formed on the viewing screen.

Figure 6 shows some of the ways in which two waves can combine at the screen. In Figure 6a, the two waves, which leave the two slits in phase, strike the screen at the central point P. Because both waves travel the same distance, they arrive at P in phase. As a result, constructive interference occurs at this location, and a bright fringe is observed. In Figure 6b, the two waves also start in phase, but in this case the upper wave has to travel one wavelength farther than the lower wave to reach point Q. Because the upper wave falls behind the lower one by exactly one wavelength, they still arrive in phase at Q, and so a second bright fringe appears at this location. At point R in Figure 6c, however, between points P and Q, the upper wave has fallen half a wavelength behind the lower wave. This means that a trough of the lower wave overlaps a crest of the upper wave; this gives rise to destructive interference at point *R*. For this reason, a dark fringe is observed at this location.

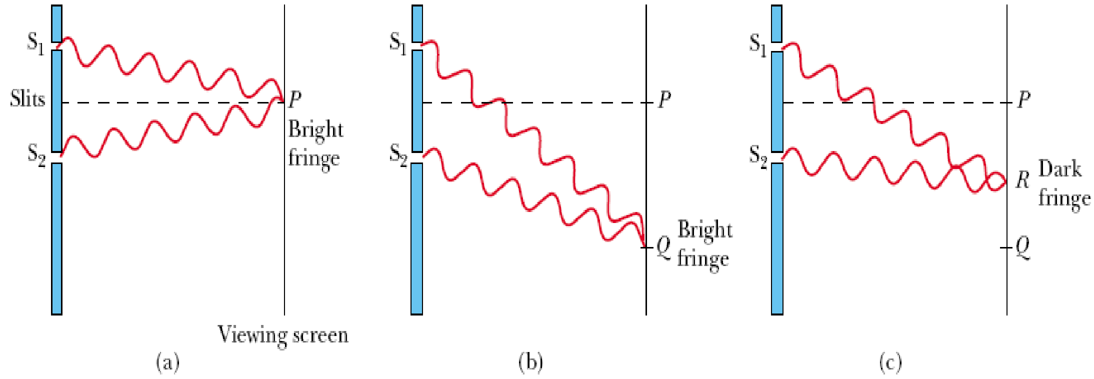

Figure 6 (a) Constructive interference occurs at point *P* when the waves combine. (b) Constructive interference also occurs at point *Q*. (c) Destructive interference occurs at *R* when the two waves combine because the upper wave falls half a wavelength behind the lower wave. (All figures not to scale.)

4.4.1 **Optical Path Difference Between the Waves at P**

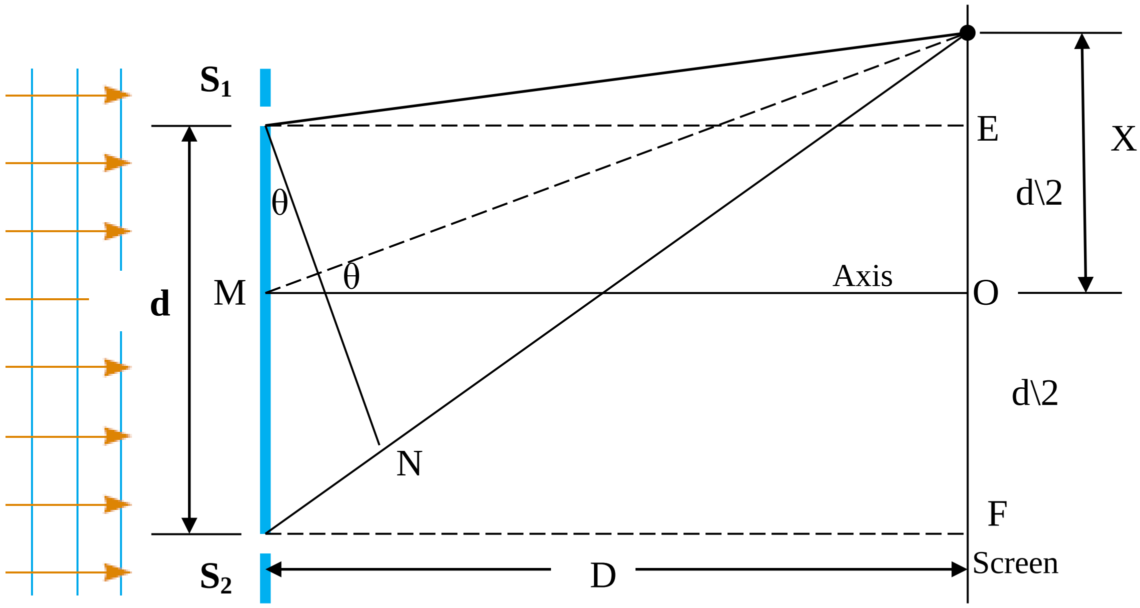

Figure 7: Geometric construction for describing Young's double-slit experiment (not to scale).

We can describe Young's experiment quantitatively with the help of Figure 7. The viewing screen is located a perpendicular distance D from the barrier containing two slits، S_1 and S_2. These slits are separated by a distance d, and the source is monochromatic. To reach any arbitrary point P in the upper half of the screen, a wave from the lower slit must travel farther than a wave from the upper slit by a distance S_2N. This distance is called the path difference ($S_2N= S_2P- S_1P$). Let the point P be at a distance X from O (figure 7). Then

PE=X - d\2 and PF=X + d\2

$$(S_2P)^2 - (S_1P)^2 = [D^2 + (PE)^2] - [D^2 + (PF)^2]$$

$$(S_2P)^2 - (S_1P)^2 = \left[D^2 + \left(X + \frac{d}{2}\right)^2\right] - \left[D^2 + \left(X - \frac{d}{2}\right)^2\right]$$

$$= \left[D^2 + \left(X^2 + 2X\frac{d}{2} + \frac{d^2}{4}\right)\right] - \left[D^2 + \left(X^2 - 2X\frac{d}{2} + \frac{d^2}{4}\right)\right]$$

$$(S_2P)^2 - (S_1P)^2 = 2Xd$$

$$(S_2P - S_1P).(S_2P + S_1P) = 2Xd$$

$$(S_2P - S_1P) = \frac{2Xd}{(S_2P + S_1P)}$$

We can approximate that $S_2P \cong S_1P \cong D$

$$\therefore \text{ path difference } = S_2P - S_1P = \frac{Xd}{D} \qquad (4.21)$$

We now find out the conditions for observing bright and dark fringes on the screen.

4.4.2 Bright Fringes

Bright fringes occur whenever the wave from S_1 and S_2 interfere constructively. The first time this occurs is at O, the axial point. There, the waves from S_1 and S_2 travel the same optical path length to O and arrive in phase. The next bright fringe occurs when the wave from S_2 travels one complete wavelength further the wave from S_1. In general constructive interference occurs if S_1P and S_2P differ by a whole number of wavelengths.

The condition for finding a bright fringe at P is that

$$S_2P - S_1P = m\lambda$$

Using equation (4.21), it means that

$$\frac{Xd}{D} = m\lambda \qquad (4.22)$$

Where m is called the ***order of fringe***.

Where m = 0 called ***zero- order of fringe*** , m = 1 called ***first- order bright fringe*** and m = 2 called ***second- order bright fringe***.

4.4.3 Dark Fringes

The first dark fringe occurs when $S_2P - S_1P$ is equal $\lambda\backslash 2$. The waves are now in opposite phase at P. The second dark fringe occurs when $S_2P - S_1P$ is equal $3\lambda\backslash 2$. The m^{th} dark fringe occurs when

$$S_2P - S_1P = (2m + 1)\lambda\backslash 2$$

The condition for finding a dark fringe is

$$\frac{Xd}{D} = (2m + 1)\frac{\lambda}{2} \qquad (4.23)$$

Where m = 0 called ***first- order dark fringe***, m = 1 called ***second- order dark fringe***.

4.4.4 Separation Between Neighboring Bright Fringe

The m^{th} order fringe occurs when

$$X_m = \frac{m\lambda D}{d}$$

And the $(m+1)^{th}$ order fringe when

$$X_{m+1} = \frac{(m+1)\lambda D}{d}$$

The fringe separation, β is given by

$$\beta = X_{m+1} - X_m = \frac{\lambda D}{d} \qquad (4.24)$$

The same result will be obtained for dark fringes. Thus, the distance between any two consecutive bright or dark fringes known as the ***fringe width*** and is same everywhere on the screen. Further, the width of the bright fringe is equal to the dark fringe. Therefore, the alternate bright and dark fringes are ***parallel***.

From equation (4.24), we find the following:

1) The fringe width β is independent of the order of the fringe. It is ***directly proportional*** to the wavelength of light, i.e. $\beta \propto \lambda$. The fringes produced by red light are less closer compared to those produced by blue light.
2) The width of the fringe is ***directly proportional*** to the distance the of the screen from the two slits $\beta \propto D$. the farther the screen, the wider is the fringe separation.
3) The width of the fringe is ***inversely proportional*** to the distance between two slits. The closer are the slits, the width will be the fringes.

4.5 Coherence

1. ***<u>Coherence time</u>: It is the average time during which the wave remain sinusoidal and phase of the wave packet can be predicted reliably.***
2. ***<u>Coherence length</u>: It is the length of the wave packet over which it may be assumed to be sinusoidal and has predictable phase.***

4.6 Conditions for Interference

We may now summarize the conditions that are to be fulfilled in order to observe a distinct well-defined interference pattern.

A) Conditions for sustained interference:

1) ***The waves from the two sources must be of the same frequency.***

If the light waves differ in frequency, the phase difference fluctuates irregularly with time. Consequently, the intensity at any point fluctuates with time and we will not observe steady interference.

2) ***The two light waves must be coherent.***

If the light waves are coherent, then they maintain a fixed phase difference over a time and space. Hence, a stationary interference pattern will be observe.

3) ***The path difference between the overlapping waves must be less than the coherence length of the waves.***

We have already learn that light is emitted in the form of wave trains and a finite coherence length characterized them. If we consider two interfering wave trains, having constant phase difference, as in figure 8, the interference effects occur due to parts QR of wave 1 and ST of wave 2.

For the parts PQ and TU interference will not occur. Therefore, the interference pattern does not appear distinctly. When the entire wave train PR overlaps on the wave train SU, interference pattern will be distinct. On the other hand, when the path difference between the waves 1 and 2 becomes very large, the wave trains arrive at different times and do not overlap on each other. Therefore, in such cases interference does not take place. The interference pattern completely vanishes if the path difference is equal to the coherence length. It is hence required that

$$\Delta < I_{coh} \tag{4.25}$$

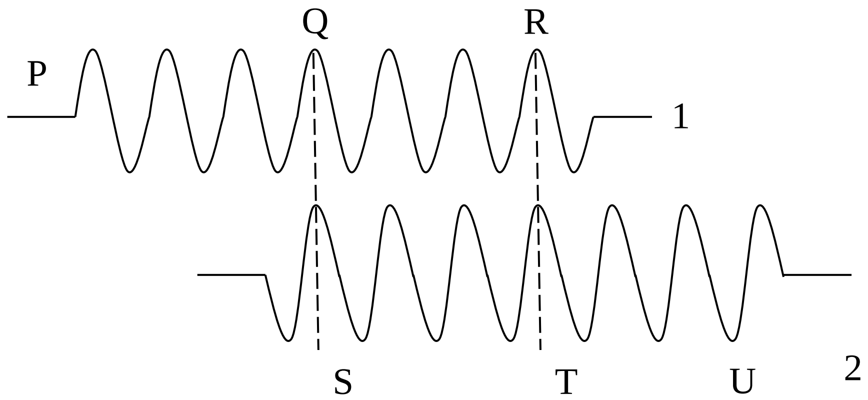

Figure 8: Partial overlap of the wave trains.

4) ***If the two sets of waves are plane polarized, their planes of polarization must be same.*** Waves polarized in perpendicular cannot produce interference effect.

B) Conditions for formation of distinct fringe pattern:

1) ***The two coherence sources must lie close to each other in order to discern the fringe pattern.*** If the sources are far apart, the fringe width will be very small and fringes are not see separately.
2) ***The distance of the screen from the two sources must be large.***
3) ***The vector sum of the overlapping electric field vectors should be zero in the dark regions*** for obtaining distinct bright and dark fringes. The sum will be zero only if the vectors are anti-parallel and have the same magnetic.

4.7 Techniques of Obtaining Interference

The techniques used for creating coherent sources of light can be divided into the following two broad classes.

1) ***Wave front splitting:*** one of the method consists in dividing a light wave front, emerging from a narrow slit, by passing it through two slits closely spaced side by side. The two parts of the same wave front travel through different path and reunite on a screen to produce fringe pattern. This is known as ***interference due to division of wave front***. This method is useful only with ***narrow sources***. Young's double slit, Fresnel's double mirror, Fresnel's biprism, Lloyd's mirror, etc employ this technique.
2) ***Amplitude splitting:*** Alternately, the amplitude (intensity) of a light wave is divided into two part, namely reflected and transmitted components, by partial reflection at a surface. The two part travel through different path and reunite to produce interference fringes. this is known as ***interference due to division of amplitude***. Optical elements such as beam splitters, mirror are used for achieving amplitude division. Interference in thin films (wedge, Newton's ring etc), Michelson's interferometer etc interferometers utilize this method. This method requires ***extended source***.

4.8 Fresnel double prism

The Fresnel double prism or biprism consists of two thin prisms joined at their bases, as shown in Figure 9. A single cylindrical wave front impinges on both prisms. The top portion of the wave front is refracted downward, and the lower segment is refracted upward. In the region of superposition, interference occurs. Here, again, two virtual sources Si and S2 exist, separated by a distance a, which can be expressed in terms of the prism angle, where D >> d. The expression for the separation of the fringes is the same as before.

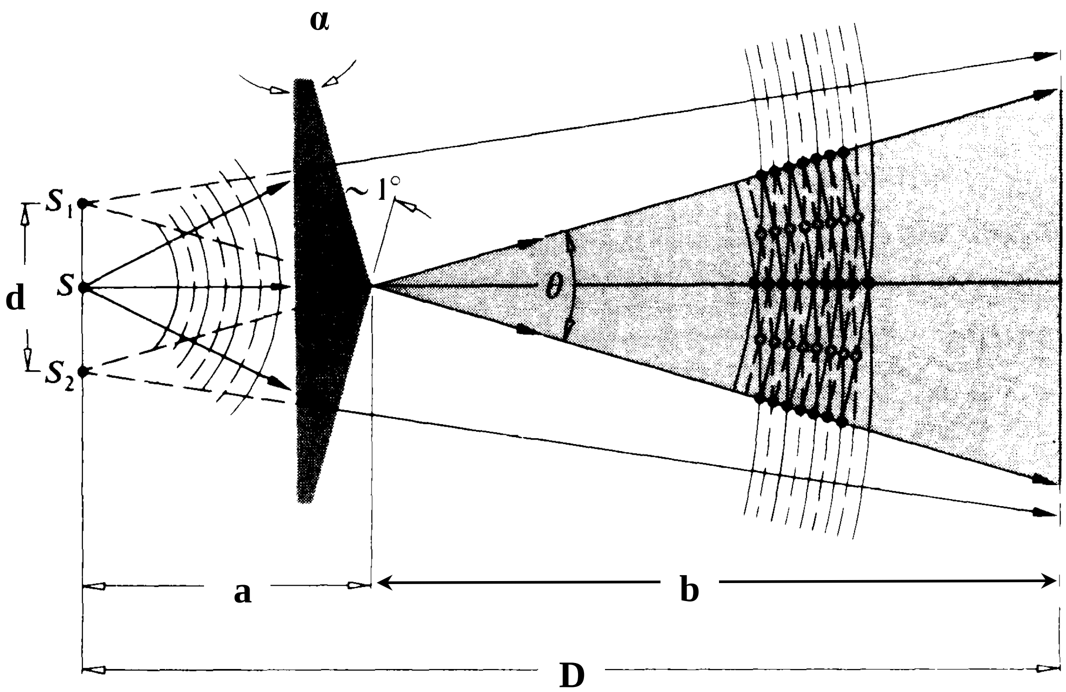

Figure 9

The width of the dark or bright fringe is given by equation (4.24)

$$\beta = \frac{\lambda D}{d}$$

where (D = a + b) is distance of the sources from the eyepiece.

4.8.1 Determination of wavelength of light

The wavelength of the light can be determined using equation (4.24). For using the relation, the values of β, D and d are to be measured. These measurements are done as follows.

1) ***Determination of fringe width β:*** When the fringes are observed in the field view of the eyepiece, the vertical cross-wire is made to coincide with the centre of one of the bright fringes. the position of the eyepiece is read on the scale, say $x_\circ$. The micrometer screw of the eyepiece is moved slowly and the number of the bright fringes N that pass across the cross-wire is counted. The position of the cross-wire is again read, say x_N. The fringe width is then given by

$$\beta = \frac{x_N - x_\circ}{N}$$

2) ***Determination of (d):*** The value of d can be determined as follows. The deviation δ produced in the path of a ray by a thin prism is given by

$$\delta = (n-1)\alpha$$

where α is the refracting angle of the prism. For the figure 9, it is seen $\delta = \theta/2$. Since d is very small, we can also write $d = a\,\theta$.

$$\delta = \frac{\theta}{2} = \frac{d}{2a} = (n-1)\alpha$$

$$d = 2a(n-1)\alpha \tag{4.25}$$

4.8.2 Interference Fringes with white light

In the biprism experiment if the slit is illuminated by white light the interference pattern consists of a central ***white fringe*** flanked on its both sides by a few coloured fringes and general illumination beyond the fringes. The central white fringe is the ***zero-order fringe***.

With monochromatic light all the bright fringes are of the same colour and it is not possible to locate zero- order fringe. Therefore, in order to locate the zero-order fringe the biprism is to be illuminated by white light.

4.8.3 Lateral Displacement of Fringes

The biprism experiment can be used to determine the thickness of given thin sheet of transparent material such as glass or mica. If a thin transparent sheet is introduced in the path of one of the two interfering beams, the fringe system gets displaced towards the beam in whose path the sheet is introduced. By measuring of displacement, the thickness of the sheet can be determined.

Suppose S_1 and S_2 are the virtual coherent monochromatic sources. The point O is equidistant from S_1 and S_2, where we obtain the ***central bright fringe***. Therefore, the optical path $S_1O= S_2O$. Let a transparent plate G of thickness t and refractive index n be introduced in the path of one of the beams(see figure 10). The optical path lengths S_1O and S_2O are now not equal and the ***central bright fringe shits*** to P from O. The light waves from S_1 to P travel partly in air and partly in the sheet G: the distance travelled in air is (S_1P – t) and that in the sheet is t.

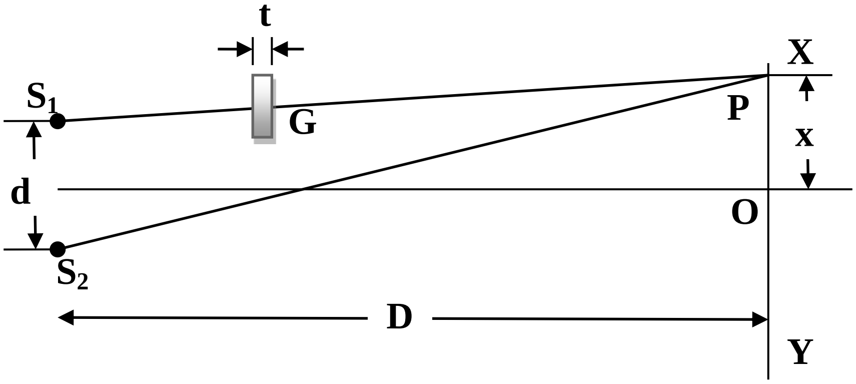

Figure 10

The optical path $\Delta_{S_1P} = (S_1P - t) + nt = S_1P + (n-1)t$

The optical path $\Delta_{S_2P} = S_2P$

The optical path difference at P is $\Delta_{S_1P} - \Delta_{S_2P} = 0$, since in the presence of the thin sheet, the optical path lengths S_1P and S_2P are ***equal*** and central zero fringe is obtained at P.

$\Delta_{S_1P} = \Delta_{S_2P}$

$[S_1P + (n-1)t] = S_2P$

$S_2P - S_1P = (n-1)t$

But according to the relation (4.21), $S_2P - S_1P = \frac{Xd}{D}$

Where x is the ***lateral shift*** of the central fringe due to the introduction of the thin sheet.

$$(n-1)t = \frac{Xd}{D}$$

Hence, the thickness of the sheet is

$$t = \frac{Xd}{D(n-1)} \tag{4.26}$$

4.9 Lloyd's Single Mirror

The last wave front-splitting interferometer that we will consider is ***Lloyd's mirror***, shown in Figure 11. It consists of a flat piece of either dielectric or metal that serves as a mirror, from which is reflected a portion of the cylindrical wave front coming from slit S. Another portion of the wave front proceeds directly from the slit to the screen. For the separation d, between the two coherent sources, we take the distance between the actual slit and its image S_1 in the mirror. The spacing of the fringes is once again given by (D/d)λ.

4.9.1 Determination of Wavelength

The fringe width is given by equation (4.24). Thus,

$$\beta = \frac{\lambda D}{d}$$

Measuring β, D and d, the wavelength can be determined .

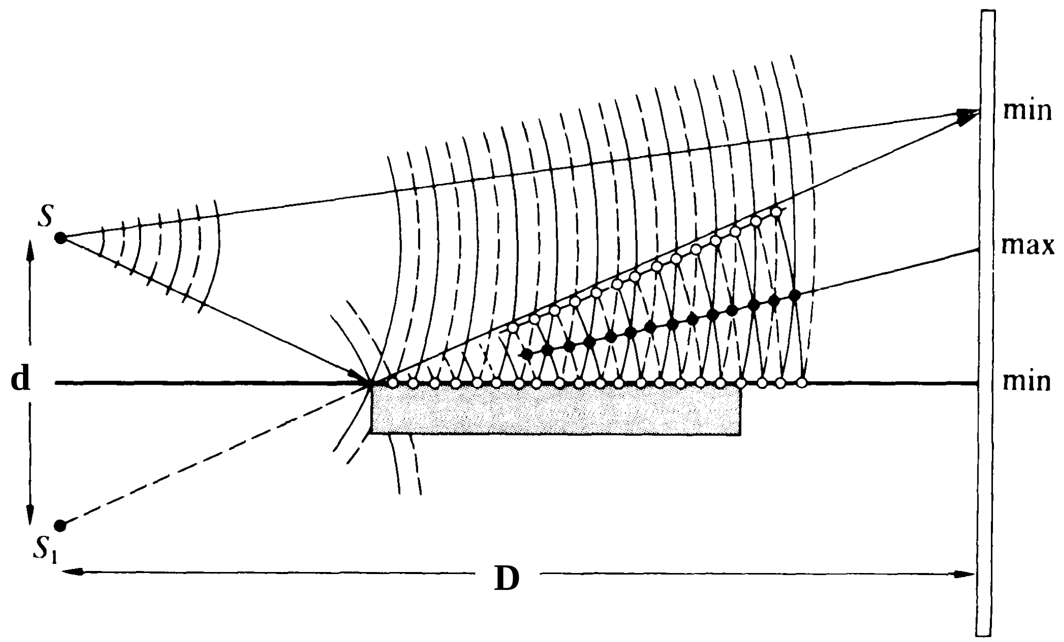

Figure 11 Lloyd's single mirror.

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) The 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.

4.10 Fresnel’s Double Mirror